Novel Panel Cointegration Tests Emending for Cross-Section Dependence with N Fixed

Hadri, K., Kurozumi, E., & Rao, Y. (2015). Novel Panel Cointegration Tests Emending for Cross-Section Dependence with N Fixed. Econometrics Journal, 18(3), 363-411. DOI: 10.1111/ectj.12054

Published in:
Econometrics Journal

Document Version:
Peer reviewed version

Novel Panel Cointegration Tests Emending for Cross-Section Dependence with N Fixed[1][2]

Kaddour Hadri
Management School
Queen's University Belfast

Eiji Kurozumi
Department of Economics
Hitotsubashi University

Yao Rao
Management School
The University of Liverpool

April 23, 2014

Revised: December 8, 2014

Revised: April 20, 2015

Revised: July 1, 2015

Abstract

In this paper, we propose new cointegration tests for single equations and panels. In both cases, the asymptotic distributions of the tests, which are derived with N fixed and $T \to \infty$, are shown to be standard normals. The effects of serial correlation and cross-sectional dependence are mopped out via long-run variances. An effective bias correction is derived which is shown to work well in finite samples; particularly when N is smaller than T. Our panel tests are robust to possible cointegration across units.

JEL classification: C12, C15, C22, C23.

Keywords: cointegration, panel cointegration, cross-section dependence, bias correction, DOLS, FCLT.

[1]Correspondence: Eiji Kurozumi, Department of Economics, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo 186-8601, Japan; e-mail: kurozumi@stat.hit-u.ac.jp.

[2]Kurozumi's research was supported by JSPS KAKENHI Grant Numbers 23243038, 24243031 and 25285067.

1. Introduction

Testing for unit roots and cointegration has been in a certain sense ignited by the seminal work of Granger and Newbold (1974) on possible spurious regressions amongst nonstationary variables, followed, some years latter, by another milestone by Engle and Granger (1987) on possible cointegration and meaningful long-term relationships amongst nonstationary variables. Since then, the concept of cointegration has played a central role in investigating long-run relationships among macroeconomic variables leading to an explosion of the literature on unit root and cointegration. The time series approach has been extended to panel data. Panel data unit root tests appeared in the early 90's whereas panel cointegration tests have been investigated in the literature since late 90's. The transposition of these tests from univariate time series to panel data was mainly motivated by the desire to increase the power of these tests by exploiting the cross-sectional dimension.

In this paper, we propose a new univariate time series cointegration test which we extend to panel data to increase its power. Therefore, in the sequel we shall deal mainly with panel data cointegration tests. However, because many statistical issues are shared between panel unit root and panel cointegration tests, we shall refer to the former whenever we deem the reference to its literature appropriate. Early contributions in the panel cointegration tests literature, are given by Kao (1999) and Pedroni (1999, 2004), among others, in which individuals are assumed to be independent which is a characteristic of the so-called first-generation panel tests. It was rapidly recognized that cross-sectional independence is an unrealistic assumption for most panel data encountered in practice. In particular, it has been shown, via simulation, that the first-generation panel unit root tests exhibit large size distortions in the presence of cross-sectional dependence (CSD) by O'Connell (1998). This is also the case for panel cointegration tests. A succession of second-generation panel cointegration tests accounting for cross-sectional dependence have been proposed in the literature. The main techniques employed to account for cross-sectional dependence are: (1) bootstrapping which generally accommodates cross-sectional dependence of general form, (2) common factors based approaches which very often requires the estimation of the number of factors and the factor loadings (3) non-linear instrumental variables, (4) rank tests which is the most recent method, and finally (5) long-run variance which could be considered as a nonparametric approach. The use of bootstrapping methods was pioneered by Maddala and Wu (1999) followed by many who applied the same technique but all of them did not establish the validity of their bootstrap. The exception, in the panel unit root tests literature, are Chang (2004) employing sieve bootstrap and Palm, Smeekes and

Urbain (2011) using block bootstrap. In both cases they established the asymptotic validity of their bootstrap tests. The advantage of the bootstrap is that it accommodates cross-sectional and temporal dependencies of general form without the need to model these dependencies and hence avoiding any possibility of misspecification. The downsides are that these methods are computationally intensive and uninformative on the structure of the cross-sectional and time series dependencies. The second approach to account for cross-sectional dependence is based on the use of common factors. In this case a specific form of the cross-sectional dependence is assumed and the way common factors are dealt with differs from one test to another. The most noticeable tests in panel unit root literature are Bai and Ng (2004), Moon and Perron (2004), Pesaran (2007) and Hadri and Kurozumi (2008, 2012). The latter consider testing the null hypothesis of stationarity against the alternative hypothesis of a unit root. Whereas the panel cointegration test correcting for CSD employing common factors were proposed *inter alia* by Gengenbach, Palm and Urbain (2006), Westerlund (2008), Westerlund and Edgerton (2008) and Bai and Carrion-i-Silvestre (2013)[3][4]. An extensive Monte-Carlo comparison of these tests can be found in Gengenbach *et al.* (2010) and De Silva, Hadri, and Tremayne (2009). Breitung and Das (2008) provides an analytical comparison of several first and second generation tests in the presence of factor structure. The advantage of this method of accounting for cross-sectional correlation is that, for small numbers of factors, it reduces the dimensionality of the covariance matrix and the number of parameters required to be estimated, which are the problems faced by other parametric methods. Further, the method allows each factor to have a unique (and possibly no) effect on each cross-section. The shortcomings of the common factors approach are that the consistent estimation of the number of factors and the loadings require that N and $T \to \infty$. The requirement that N should go to infinity is especially problematic in the sense that it puts a limit on the practical applicability of the factor based tests in macroeconomics and finance where N is typically relatively small. Another concern is the problem of possible misspecification of the factor structure which can result in severe size distortions (cf. Breitung and Das, 2008). A third method amending for CSD via non-linear IV methodology was pioneered by Chang (2002) and improved by Chang and Song (2009) for testing panel unit root. The same approach was applied by Chang and Nguyen (2012) in the context of panel cointegration tests. This is a very innovative and appealing approach to deal with some issues of testing panel unit root and cointegration but did not seem to receive enough attention so far. The

[3] Baltagi (2008) and Breitung and Pesaran (2008) provide comprehensive surveys on the theoretical advantage of using panel data.

[4] Single equation tests for the null of cointegration in the literature include Hansen (1992), Quintos and Phillips (1993), Shin (1994), Jansson (2005) and Kurozumi and Arai (2008).

fourth approach is the most recent one. It is based on the rank of the long-run variance matrix of the N-dimensional vector of stacked observations of the observed panel data, where N denotes the cross-section sample size. This makes the tests suitable both as conventional panel unit root tests with the corresponding null and alternative hypotheses, or, more generally, as flexible rank tests that allow one to determine the number of common trends in the panel. These tests are proposed by Pedroni, Vogelsang, Wagner and Westerlund (2015), thereafter PVWW and can accommodate very general forms of both serial and cross-sectional dependence. Finally, the fifth approach to adjust for CSD has been offered for the first time by Driscoll and Kraay (1998) in panel GMM setting. It is a nonparametric approach based on the construction of a cross-sectional consistent estimator similar to the Newey and West (1987) time series estimator. Similar idea is employed in this paper to emend for arbitrary cross-sectional dependence and serial correlation in a non-parametric way hence, avoiding any potential misspecification of these dependencies.

Another issue uncovered by Banerjee, Marcellino and Osbat (2004) through simulations is that panel unit root and cointegration tests have severely distorted size in presence of cross units cointegration. Some panel tests have been corrected for this including Chang and Song (2009), Chang and Nguyen (2012), HLM (2005), Palm, Smeekes and Urbain (2011), PVWW (2015) and the present paper.

A third issue in panels of the type we are considering here, is the asymptotic theory to be used. The limit theory for this class of panel data has been developed in a seminal paper by Phillips and Moon (1999). In their paper, they study *inter alia* the limit theory that allows for both sequential limits, wherein $T \to \infty$ followed by $N \to \infty$, and joint limits where $T, N \to \infty$ simultaneously. They also mention, in the same paper, the diagonal path limit theory in which the passage to infinity is done along a specific diagonal path. The drawback of sequential limits is that in certain cases, they can give asymptotic results which are misleading. The downside of diagonal path limit theory is that the assumed expansion path $(T(N), N) \to \infty$ may not provide an appropriate approximation for a given (T, N) situation. Finally, the joint limit theory requires, generally, a rate condition on the relative speed of T and N going to infinity. In this paper, we consider a limit theory in which N is fixed and T is allowed to go to infinity. In this case the asymptotics are far easier to derive than in the case of the joint asymptotic where $T, N \to \infty$ simultaneously. Because our asymptotic results do not depend on $N \to \infty$, these tests are conveniently suited for typical macroeconomic and financial applications where N is usually relatively smaller than T. The fixed N asymptotics have been considered by Chang (2004), Chortareas and Kapetanios (2009), Hanck (2009), Palm, Smeekes and Urbain (2011), PVWW (2015) and Moon and Perron (2012). This asymptotic theory cannot be used in the factor model context which requires that N and $T \to \infty$ in order to estimate consistently the number of factors and the loadings. This puts a limit on the practical applicability of the factor based tests in macroeconomic and finance panels where N is typically relatively small.

An additional aspect in panel data we are concerned with in this paper is the way the null hypothesis and the alternative are expressed. In testing panel unit root most of the tests have as the null hypothesis that all the units are jointly nonstationary against the alternative that some of the units are stationary. In all these studies the unit root is the null hypothesis to be tested, and it is a well-known fact that the way in which classical hypothesis testing is carried out ensures that the null hypothesis is accepted unless there is strong evidence to the contrary. This is confirmed in the time series literature by the fact that it has been found that standard unit root tests are not very powerful against relevant alternatives and fail to reject the null hypothesis for many economic series. These studies suggest that, in trying to decide by classical methods whether economic data are stationary or integrated, it would be useful to perform tests of the null hypothesis of stationarity as well as tests of the null hypothesis of a unit root. Only few panels consider the null hypothesis of stationarity these include HLM (2005), Hadri and Rao (2008) and Hadri and Kurozumi (2012). In the panel cointegration literature most tests assume the null hypothesis of no panel cointegration against the alternative of cointegration. Hence, rejection of the null hypothesis is often understood as the existence of (partial) panel cointegration. However, from the view of classical hypothesis testing, if we are primary concerned about cointegration, it seems more natural to choose panel cointegration as the null hypothesis. In addition, panel cointegration (no panel cointegration) would be strongly supported if the null of panel cointegration is accepted (rejected) while the null of no panel cointegration is rejected (accepted) by some tests. Therefore, we may see that tests for the null of panel cointegration complement no panel cointegration tests. The panel test we are proposing in this paper is one of the rare which has as the null hypothesis of joint cointegration.

A final issue with testing for unit root in panels is what happens when the null of non-stationarity which is a joint hypothesis is rejected. As a consequence the null hypothesis of a unit root may be rejected even if only one of the unit is stationary. Thus, the possibility emerges that small groups of cross-sectional units in the panel, that share particular features, may drive the results. Therefore, panel unit root tests are sensitive to the selection of series included in the panel. In particular if one rejects the joint unit root hypothesis, one cannot know which series caused the rejection. Chortareas and Kapetanios (2008) proposes the Sequential Panel Selection Method (SPSM) which consists of carrying out a sequence of panel unit root tests on panels of decreasing size. After a rejection, a researcher removes from the panel the series with the most evidence in favour of stationarity. One then continues until the joint test of a unit root for the remaining series in the panel is no longer rejected. A different approach was suggested by Ng (2008) who estimates the fraction of nonstationary series. She conjectures that one can then identify the I(1) and I(0) series by ordering them according to the magnitude of their autoregressive parameter. Hanck (2009) procedure seeks to control for the Familywise Error Rate (FWER) when performing multiple hypotheses tests in panels. The bootstrap is used to estimate one of the statistics in order to ensure a joint asymptotic coverage probability $1 - \alpha$. Moon and Perron (2012) use the false discovery (FDR) to uncover which series is I(1) and which series is stationary. FDR controlling procedures exert a less stringent control over false discovery compared to familywise error rate (FWER) procedures (such as the Bonferroni correction), which seeks to reduce the probability of even one false discovery, as opposed to the expected proportion of false discoveries. Thus FDR procedures have greater power at the cost of increased rates of type I errors, i.e., rejecting the null hypothesis of no effect when it should fail to be rejected. These methods can be adapted to the panel cointegration tests when the null joint hypothesis of cointegration or non-cointegration is rejected. However, we shall not pursue this important issue here as it needs a paper on its own and therefore leave it for possible future research.

As panel unit root tests, panel cointegration tests are very popular amongst empirical researchers. The extension to panel cointegration makes it possible to investigate international, regional or industrial relations. Banerjee and Wagner (2009) investigated growth convergence. The purchasing power parity (PPP) hypothesis has often been examined in the literature using panel unit roots/cointegration techniques. See, for example, Frankel and Rose (1996), Papell (1997), Wagner (2008a) and Hanck (2009) among others. There are many other empirical investigations employing panel cointegration tests including, the exchange rate pass-through by Banerjee and Carrion-i-Silvestre (2013), the environmental Kuznets curve by Wagner (2008b), the Fisher effect by Westerlund (2008) and the bond market by Matsuki (2015).

As noted above, in this paper we propose a new test to test for cointegration in univariate time series but more importantly we shall transpose our test to panels to increase its power. Our panel cointegration test accommodate cross-sectional dependence of arbitrary form and treat the possible serial correlation non-parametrically hence avoiding any possible misspecification. Our test is invariant to the possible presence of cross-sectional cointegration. Our asymptotic theory does not require $N \to \infty$ which makes our tests suitably appropriate for typical macroeconomic and financial applications where N is generally smaller than T. Our statistic is simple to construct and conveniently has a limiting distribution under the null hypothesis that is standard normal and therefore there is no need to compute bootstrap critical values. Another upside of our test is that the null hypothesis is cointegration which is the appropriate null when we wish to test for a long-run relationships amongst non-stationary economic or financial variables. More precisely, our tests are based on the autocovariances of the error term as considered by Harris, Leybourne and McCabe (2005, hereafter HLM) and our test statistics are shown to be asymptotically free of nuisance parameters. As a result, we can rely on the asymptotic critical values to test for panel cointegration. In addition, we propose the bias corrected version of the autocovariance based test to improve the finite sample properties, because, as discussed in HLM (2005), the test statistic based on the autocovariances has a negative bias in finite samples, which makes the test conservative. We will show by Monte Carlo simulations that our bias correction works very well to control the empirical size of the proposed test.

The layout of this paper is as follows. In Section 2, we review the autocovariance based test proposed by Harris, McCabe and Leybourne (2003) and HLM (2005). The new univariate cointegration tests are analyzed in the following section. Section 4 investigates the novel panel cointegration tests. The finite sample property of our tests are investigated in Section 5. Finally, Section 6 offers some concluding remarks, and all proofs are collected in the Appendix.

2. Review of the Autocovariance Based Test

In this section, we briefly review stationarity tests based on the autocovariance proposed by Harris, McCabe and Leybourne (2003, hereafter HML) and Harris, Leybourne and McCabe (2005, hereafter HLM). Let us consider the following local level model[5]:

$$y_t = \mu + z_t \quad \text{for} \quad t = 1, 2, \cdots, T,$$

and suppose that we want to test for the null hypothesis that z_t is stationary whereas it is a unit root process under the alternative. HML (2003) note the differences in the convergence order of the sample

[5]HML (2003) and HLM (2005) allowed for deterministic regressors in addition to a constant but we restrict our attention to the local level model in order to simplify the explanation.

autocovariance under the null and the alternative hypotheses,

$$\frac{1}{T-K}\sum_{t=K+1}^{T}\hat{z}_t\hat{z}_{t-K} \xrightarrow{p} E[(y_t-\mu)(y_{t-K}-\mu_K)] \equiv C_K \quad \text{under the null hypothesis}$$

$$\frac{1}{(T-K)^2}\sum_{t=K+1}^{T}\hat{z}_t\hat{z}_{t-K} \xrightarrow{d} \int_0^1 \tilde{B}^2(r)dr \quad \text{under the alternative}$$

for a given lag order K, where $\hat{z}_t = y_t - \bar{y}$ and $\tilde{B}(r)$ is a demeaned Brownian motion. Although it seems inconvenient to use the sample autocovariance as a test statistic because it converges to a fixed value C_K, HML (2003) notice that $C_K \to 0$ as $K \to \infty$ and thus the central limit theorem (CLT) for the sample autocovariance with a suitable normalization is expected to hold as K goes to infinity. In fact, they showed that

$$\frac{\frac{1}{\sqrt{T-K}}\sum_{t=K+1}^{T}\hat{z}_t\hat{z}_{t-K}}{\hat{\omega}_{zz}} \xrightarrow{d} N(0,1) \quad \text{under the null hypothesis,} \tag{1}$$

where $\hat{\omega}_{zz}^2$ is the kernel estimator of the long-run variance based on $\hat{z}_t\hat{z}_{t-K}$, whereas the left-hand side diverges to infinity under the alternative. They also proposed a test for heteroskedastic cointegration using a similar principle.

The above stationarity test based on the autocovariance was extended to a panel stationarity test by HLM (2005). For a panel data model given by

$$y_{i,t} = \mu_i + z_{i,t} \quad \text{for} \quad i = 1, 2, \cdots, N \quad \text{and} \quad t = 1, 2, \cdots, T,$$

we have the regression residuals normalized by the standard deviation; that is,

$$\tilde{z}_{i,t} = \frac{\hat{z}_{i,t}}{\hat{\sigma}_{i,z}}, \quad \text{where } \hat{z}_{i,t} = z_{i,t} - \bar{z}_i \text{ and } \hat{\sigma}_{i,z} \text{ is the sample standard deviation of } \hat{z}_{i,t}.$$

Then, the test statistic for panel stationarity is constructed by pooling the sample autocovariances across cross-sections, which is given by

$$\hat{S}_K = \frac{\tilde{C}_K}{\hat{\omega}_a}, \quad \text{where} \quad \tilde{C}_K = \frac{1}{\sqrt{T-K}}\sum_{t=K+1}^{T}\tilde{a}_{K,t} \quad \text{with} \quad \tilde{a}_{K,t} = \sum_{i=1}^{N}\tilde{z}_{i,t}\tilde{z}_{i,t-K}$$

and $\hat{\omega}_a^2$ is the long-run variance estimator based on $\tilde{a}_{K,t}$. HLM (2005) showed that $\hat{S}_K \xrightarrow{d} N(0,1)$ under the null hypothesis whereas it diverges to infinity under the alternative.

Although the size of the above test can be controlled at least asymptotically, HLM (2005) showed that $\hat{S}_K$ suffers from under-size distortion in finite samples because of the negative bias of the test statistic.

Since $\hat{z}_{i,t} = z_{i,t} - \bar{z}_i$, we can see that

$$\frac{1}{\sqrt{T-K}} \sum_{t=K+1}^{T} \tilde{z}_{i,t}\tilde{z}_{i,t-K} = \frac{1}{\hat{\sigma}_{i,z}^2 \sqrt{T-K}} \sum_{t=K+1}^{T} z_{i,t} z_{i,t-K} - \frac{1}{\hat{\sigma}_{i,z}^2 \sqrt{T-K}} \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} z_{i,t} \right)^2 + o_p\left(\frac{1}{\sqrt{T}}\right),$$

and thus the negative bias comes from the second term on the right-hand side of the above equation. Note that this negative bias accumulates when we pool the sample autocovariances, so that the panel stationarity test tends to be severely undersized as N gets larger. Because the expectation of $(T^{-1/2}\sum_{t=1}^{T} z_{i,t})^2$ is approximated by the long-run variance of its limiting distribution, HLM (2005) proposed the following bias-corrected version of the test statistic:

$$\tilde{S}_K = \frac{\tilde{C}_K + \tilde{b}}{\hat{\omega}_a} \quad \text{where} \quad \tilde{b} = \frac{1}{\sqrt{T-K}} \sum_{i=1}^{N} \frac{\hat{\omega}_{i,z}^2}{\hat{\sigma}_{i,z}^2}$$

with $\hat{\omega}_{i,z}^2$ being the long-run variance estimator based on $\hat{z}_{i,t}$. Because the bias term is negligible when T is large, we still have $\tilde{S}_K \xrightarrow{d} N(0,1)$ under the null hypothesis.

3. Univariate Cointegration Test

3.1. Model and assumptions

We start with a univariate cointegrating regression model given by

$$y_t = \beta' X_t + u_t \quad \text{for} \quad t = 1, 2, \cdots, T, \tag{2}$$

where $X_t = [1, x_t']'$ (constant case) or $X_t = [1, t, x_t']'$ (trend case), y_t and x_t are 1-and p_x-dimensional processes with

$$x_t = x_{t-1} + v_t \quad \text{and} \quad u_t = \rho u_{t-1} + u_t^*.$$

We make the following assumption for u_t^* and v_t:

Assumption 1 *(a) $[u_t^*, v_t']'$ is a vector linear process given by*

$$\begin{bmatrix} u_t^* \\ v_t \end{bmatrix} = \sum_{j=0}^{\infty} \Phi_j \varepsilon_{t-j} \quad \text{with} \quad \sum_{j=0}^{\infty} j^2 \|\Phi_j\| < \infty,$$

where $\{\varepsilon_t\}$ is an (p_x+1)-dimensional i.i.d. sequence with mean 0 and variance given by Σ_ε, which is positive definite, and has the finite fourth order moments.
(b) The spectral density of $[u_t^*, v_t']'$, denoted by $f(\lambda) \equiv (2\pi)^{-1}\Phi(e^{-i\lambda})\Sigma_\varepsilon\Phi'(e^{i\lambda})$, is nonsingular and $f(\lambda) \geq \alpha I_{p_x+1}$ for some $\alpha > 0$ for all $\lambda \in [0, \pi]$.

Assumption 1 is standard in the literature on single cointegration tests, except for the 2-summability condition. Assumption 1(a) implies that $[u_t^*, v_t']'$ is stationary and that there is no cointegrating relation among x_t. This assumption is used for the functional central limit theorem to hold. The 2-summability of $\{\Phi_j\}$ is stronger than usual but we need this condition to derive the bias later. The assumption on the spectral density in (b) will be used to derive the leads and lags expression as considered by Saikkonen (1991). We also note that, since $\{\varepsilon_t\}$ is an *i.i.d.* sequence with the finite fourth order moments, exercise 2.13 of Brillinger (1981) implies that $[u_t^*, v_t']'$ satisfies Assumption 2.6.2 of Brillinger (1981). That is, the fourth order cumulants of $[u_t^*, v_t']'$, which are denoted by $\kappa_{ijkl}(m_1, m_2, m_3)$, satisfy

$$\sum\sum_{m_1,m_2,m_3=-\infty}^{\infty}\sum |\kappa_{ijkl}(m_1, m_2, m_3)| < \infty.$$

The testing problem we consider is given by

$$H_0: \ |\rho| < 1 \quad \text{vs.} \quad H_1: \ \rho = 1.$$

That is, y_t is cointegrated with x_t under the null hypothesis whereas they are not cointegrated under the alternative. Note that under the null hypothesis, $[u_t, v_t']'$ also satisfies the same conditions as given by Assumption 1.

Since it is known that $\tilde{D}_T(\hat{\beta}_{ols} - \beta)$ converges in distribution where $\hat{\beta}_{ols}$ is obtained by regressing y_t on X_t and $\tilde{D}_T = diag\{\sqrt{T}, T\ I_{p_x}\}$ (constant case) or $\tilde{D}_T = diag\{\sqrt{T}, T\sqrt{T}, T\ I_{p_x}\}$ (trend case), we can see that the same weak convergence holds as given by (1) with $\hat{z}_t$ replaced by $\hat{u}_t$. That is, we can test for (panel) cointegration using autocorrelations constructed from the least squares residuals, at least asymptotically. However, such a test suffers from under-size distortion as we will see in the simulation section. The reason for the under-size distortion in finite samples is similar to the case of the stationarity test. In the case of cointegration model (2), we have

$$\begin{aligned}
&\frac{1}{\sqrt{T-K}}\sum_{t=K+1}^{T}\hat{u}_t\hat{u}_{t-K} \\
&= \frac{1}{\sqrt{T-K}}\sum_{t=K+1}^{T}u_t u_{t-K} - (\hat{\beta}_{ols} - \beta)'\frac{1}{\sqrt{T-K}}\sum_{t=K+1}^{T}X_{t-K}u_t \\
&\quad -(\hat{\beta}_{ols} - \beta)'\frac{1}{\sqrt{T-K}}\sum_{t=K+1}^{T}X_t u_{t-K} + (\hat{\beta}_{ols} - \beta)'\frac{1}{\sqrt{T-K}}\sum_{t=K+1}^{T}X_t X_{t-K}'(\hat{\beta}_{ols} - \beta) \\
&= \frac{1}{\sqrt{T-K}}\sum_{t=K+1}^{T}u_t u_{t-K} - \frac{1}{\sqrt{T-K}}\sum_{t=1}^{T}u_t X_t'\left(\sum_{t=1}^{T}X_t X_t'\right)^{-1}\sum_{t=1}^{T}X_t u_t + o_p\left(\frac{1}{\sqrt{T}}\right), \qquad (3)
\end{aligned}$$

where the first term on the last expression is the leading term. Although the second term in the last expression disappears asymptotically, this term does affect the finite sample performance. Because this term is in quadratic form and then non-negative, we can see that the autocovariance based test tends to be negatively biased in finite samples due to the second term of expression (3).

One of the possible solutions for the under-size distortion is the bias correction as suggested by HLM (2005), but in the case of cointegration models, the estimation of the bias becomes much difficult because the OLS estimator is second-order biased in cointegrating regressions.

In order to eliminate the second-order bias from the OLS estimator, we exploit the dynamic ordinary least squares (DOLS) technique[6] considered by Phillips and Loretan (1991), Saikkonen (1991) and Stock and Watson (1993). As we will see in Lemma 3, we can easily estimate the bias, which will be defined later, using DOLS regressions. Under Assumption 1 and the null hypothesis, we have the following leads and lags expression by Theorem 8.3.1 of Brillinger (1981):

$$u_t = \sum_{j=-\infty}^{\infty} \pi_j' v_{t-j} + \eta_t, \tag{4}$$

where $E[v_s \eta_t] = 0$ for all s and t, and the transfer function associated with $\{\pi_j\}$ is given by $f_{uv}(\lambda) f_{vv}^{-1}(\lambda)$ with $f_{uv}(\lambda)$ and $f_{vv}(\lambda)$ being the corresponding blocks of $f(\lambda)$. Then, the assumption of the 2-summability of $\{\Phi_j\}$ implies that $\{\pi_j\}$ is also 2-summable. In addition, because $[u_t, v_t']'$ is a linear process with *i.i.d.* innovations, η_t can be expressed as

$$\eta_t = \sum_{j=-\infty}^{\infty} \phi_j \xi_{t-j} \quad \text{with} \quad \sum_{j=-\infty}^{\infty} |j|^2 |\phi_j| < \infty, \tag{5}$$

where $\{\xi_t\}$ is an independent sequence with mean 0, variance σ_ξ^2 and the finite fourth order moments. By replacing u_t in (2) with (4), we have

$$y_t = \beta' X_t + \sum_{j=-\infty}^{\infty} \pi_j' v_{t-j} + \eta_t.$$

By truncating infinite leads and lags at $j = \pm M$, we obtain the DOLS regression as follows:

$$y_t = \beta' X_t + \sum_{j=-M}^{M} \pi_j' v_{t-j} + \eta_t^*, \quad \text{for } t = M+1, \cdots, T-M, \tag{6}$$

where $\eta_t^* = \eta_t + \sum_{j>|M|} \pi_j' v_{t-j}$. Note that the truncation points can be different at the leads and the lags; in fact, the finite sample performance with the different truncation points could be better in some

[6] We also considered the fully modified (FM) regression proposed by Phillips and Hansen (1990). Although this estimator is free from the second order bias, it can be shown that the tedious bias in the covariance based test statistic still remains even if the FM method is applied and thus we do not pursue the FM technique in this paper.

cases as investigated by Hayakawa and Kurozumi (2008) and Choi and Kurozumi (2012). In this paper, the same truncation points are used only for notational convenience.

In the following, we consider constructing a test statistic based on regression (6) and thus for notational convenience, we re-define $T = T - 2M$ and denote the effective sample period $t = M + 1, \cdots, T - M$ as $t = 1, \cdots, T$.

As discussed in Saikkonen (1991), the truncation point M must diverge to infinity at a suitable rate and we make the following assumption on the divergence rate of M:

Assumption 2 *As* $T \to \infty$,

$$\frac{M^4}{T} \to 0, \tag{7}$$

$$\sqrt{T} \sum_{|j|>M} \|\pi_j\| \to 0. \tag{8}$$

Conditions (7) and (8) gives the upper and lower bounds for the divergence rate of M, respectively. Note that Saikkonen (1991) assumed $M^3/T \to 0$, which is weaker than (7) and sufficient to guarantee the asymptotic normality of π_j for a given j. The stronger assumption 2 is required in order to evaluate the bias term in our cointegration test. Note that, as shown by Kejriwal and Perron (2008), we can relax Assumption 2 as far as the efficient estimation of β is concerned.

3.2. Cointegration test with DOLS regressions

We construct the test statistic following HML (2003). Let $\hat{\eta}_t^*$ be the regression residuals from DOLS regression (6) and the standardized version[7] is given by

$$\tilde{\eta}_t^* = \frac{\hat{\eta}_t^*}{\hat{\sigma}_\eta}, \quad \text{where} \quad \hat{\sigma}_\eta^2 = \frac{1}{T}\sum_{t=1}^{T} \hat{\eta}_t^{*2}.$$

Then, the test statistic for the null of cointegration is given by

$$\hat{S}_K = \frac{\tilde{C}_K}{\hat{\omega}_a} \quad \text{where} \quad \tilde{C}_K = \frac{1}{\sqrt{T-K}} \sum_{t=K+1}^{T} \tilde{a}_{K,t} \quad \text{with} \quad \tilde{a}_{K,t} = \tilde{\eta}_t^* \tilde{\eta}_{t-K}^*,$$

and $\hat{\omega}_a^2$ is the long-run variance estimator of $\tilde{a}_{K,t}$ with the Bartlett kernel given by

$$\hat{\omega}_a^2 = \hat{\gamma}_{a,0} + 2\sum_{j=1}^{J}\left(1 - \frac{j}{J+1}\right)\hat{\gamma}_{a,j} \quad \text{where} \quad \hat{\gamma}_{a,j} = \frac{1}{T-K}\sum_{t=K+j+1}^{T} \tilde{a}_{K,t}\tilde{a}_{K,t-j} \tag{9}$$

[7] Exactly speaking, it is not necessary for the residuals to be standardized as far as the univariate case is concerned; the standardization is required only for the panel cointegration test in order for the test statistic to be scale invariant. We standardize them in the univariate case just because the univariate cointegration test can be seen as a special case of the panel cointegration test.

and J is the bandwidth of order $o(T^{1/2})$.

We would like to show that the functional central limit theorem (FCLT) holds for $\tilde{C}_K$, but we cannot directly apply theorems in HML (2003) because they assume a causal linear process for the stochastic term z_t whereas η_t in our model is not a causal but a linear process with leads and lags of the innovations $\{\xi_t\}$. Then, we first have to establish the Beveridge–Nelson (B–N) decomposition for $\eta_t\eta_{t-K}$. In the following, the coefficients and the lag polynomials depend on K but we suppress it for notational convenience.

Lemma 1 *For $\{\eta_t\}$ satisfying (5), we have*

$$\eta_t\eta_{t-K} = \sum_{j=1}^{\infty} G_j\xi_t\xi_{t-j} - \Delta\tilde{r}_t - \Delta^+\tilde{r}_t^+ + r_{1,t} + r_{2,t} + r_{3,t}, \tag{10}$$

where $\Delta = 1 - L$ and $\Delta^+ = 1 - L^{-1}$ with L being the lag operator, $G_j = G_{1,j} + G_{2,j}$ with

$$G_{1,j} = \sum_{\ell=1-(j\wedge K)}^{K-1} \phi_\ell\phi_{j+\ell-K} \text{ and } G_{2,j} = \begin{cases} \sum_{\ell=1}^{K-j-1} \phi_{\ell-K}\phi_{j+\ell}, & (j = 1, \cdots, K-2), \\ 0, & (j > K+2), \end{cases}$$

$\tilde{r}_t = \tilde{r}_{1,t} + \tilde{r}_{2,t}$ with

$$\begin{aligned}
\tilde{r}_{1,t} &= \sum_{j=1}^{\infty} \tilde{G}_{1,j}(L)\xi_t\xi_{t-j} \quad \text{where } \tilde{G}_{1,j}(L) = \sum_{\ell=0}^{K-2} \tilde{G}_{1,\ell}L^\ell \text{ with } \tilde{G}_{1,\ell} = \sum_{i=\ell+1}^{K-1} \phi_i\phi_{i+j-K}, \\
\tilde{r}_{2,t} &= \sum_{j=1}^{K-2} \tilde{G}_{2,j}(L)\xi_t\xi_{t-j} \quad \text{where } \tilde{G}_{2,j}(L) = \sum_{\ell=0}^{K-j-2} \tilde{G}_{2,\ell}L^\ell \text{ with } \tilde{G}_{2,\ell} = \sum_{i=\ell+1}^{K-j-1} \phi_{i+j}\phi_{i-K}, \\
\tilde{r}_t^+ &= \sum_{j=2}^{\infty} \tilde{G}_j^+(L)\xi_t\xi_{t-j} \quad \text{where } \tilde{G}_j^+(L) = \sum_{\ell=2-(j\wedge K)}^{0} \tilde{G}_\ell^+L^\ell \text{ with } \tilde{G}_\ell^+ = \sum_{i=1-(j\wedge K)}^{\ell-1} \phi_i\phi_{i+j-K},
\end{aligned}$$

$$r_{1t} = \sum_{j=1}^{K-1} \phi_j\phi_{j-K}\xi_{t-j}^2, \quad r_{2t} = \sum_{|j|\geq K}^{\infty} \sum_{\ell=-\infty}^{\infty} \phi_j\phi_\ell\xi_{t-j}\xi_{t-K-\ell}, \quad r_{3t} = \sum_{j=-K+1}^{K-1} \sum_{\ell=-\infty}^{-K} \phi_j\phi_\ell\xi_{t-j}\xi_{t-K-\ell}.$$

Lemma 1 implies that $\eta_t\eta_{t-K}$ can be decomposed into the first term on the right-hand side of (10) plus the remaining terms, the former of which is a martingale difference array. In order to establish the FCLT for the partial sum process of $\eta_t\eta_{t-K}$, we make the following assumption on the divergence rate of K.

Assumption 3 *The lag order K diverges to infinity at a rate of T^δ for $1/4 \leq \delta < 1$.*

The divergence rate of K is related with the establishment of Lemma A.2(ii) in the appendix, the proof of which implies that if, in general, $\{\phi_i\}$ is j-summable, then K could be T^δ for $1/(2j) \leq \delta < 1$.

Since $\{\phi_i\}$ is 2-summable in our case, we make Assumption 3. Note that Assumptions 2 and 3 imply that $M/K \to 0$, which is required in the proofs of the lemmas and theorems.

From expression (10), the FCLT for a sequence of martingale difference arrays can be applied to the first term on the right-hand side of (10) by the following Lemma 2 while the differencing operators $\Delta = 1 - L$ and $\Delta^+ = 1 - L^{-1}$ avoid from accumulating the effect of $\tilde{r}_t$ and $\tilde{r}_t^+$. Intuitively, the partial sums of the remaining terms $r_{1,t}$, $r_{2,t}$ and $r_{3,t}$ become negligible because they include ϕ_j for $j \geq K$, which converges to zero sufficiently rapidly.

Lemma 2 *Suppose that Assumptions 1 and 3 hold. Under the null hypothesis, the following FCLT holds as $T \to \infty$:*

$$\frac{1}{\sqrt{T-K}} \sum_{t=1}^{[Tr]} \eta_t \eta_{t-K} \Rightarrow B(r), \tag{11}$$

where $[a]$ is the largest integer less than a, $0 \leq r \leq 1$, $\Rightarrow$ signifies weak convergence of the associated probability measures, and $B(r)$ is a Brownian motion with the variance $\omega_a^2 \equiv \sigma_\xi^4 \lim_{K\to\infty} \sum_{j=1}^{\infty} G_j^2$.

Note that (11) holds only when $K \to \infty$ at a suitable rate; otherwise, the left-hand side apparently goes to infinity.

We are now in a position to apply Lemma 2 to the residuals in DOLS regression (6). Since

$$\hat{\eta}_t^* = \eta_t - (\hat{\beta} - \beta)' X_t - (\hat{\Pi} - \Pi)' V_t + e_t,$$

where $\hat{\beta}$ and $\hat{\Pi}$ are the estimators of β and Π in (6) with $\Pi = [\pi_M, \pi_{M-1}, \cdots, \pi_{-M}]$, $V_t = [v'_{t-M}, v'_{t-M+1}, \cdots, v'_{t+M}]'$, and $e_t = \sum_{|j|>M} \pi'_j v_{t-j}$, we have

$$\frac{1}{\sqrt{T-K}} \sum_{t=K+1}^{T} \hat{\eta}_t^* \hat{\eta}_{t-K}^* = \frac{1}{\sqrt{T-K}} \sum_{t=K+1}^{T} \eta_t \eta_{t-K} + \frac{1}{\sqrt{T-K}} \left(R_{\beta,T} + R_{\Pi,T} + R_T \right), \tag{12}$$

where

$$R_{\beta,T} = (\hat{\beta} - \beta)' \sum_{t=K+1}^{T} X_t X_{t-K} (\hat{\beta} - \beta) - (\hat{\beta} - \beta)' \sum_{t=K+1}^{T} X_{t-K} \eta_t - (\hat{\beta} - \beta)' \sum_{t=K+1}^{T} X_t \eta_{t-K}, \tag{13}$$

$$R_{\Pi,T} = (\hat{\Pi} - \Pi)' \sum_{t=K+1}^{T} V_t V_{t-K} (\hat{\Pi} - \Pi) - (\hat{\Pi} - \Pi)' \sum_{t=K+1}^{T} V_{t-K} \eta_t - (\hat{\Pi} - \Pi)' \sum_{t=K+1}^{T} V_t \eta_{t-K}, \tag{14}$$

$$\begin{aligned} R_T \;=\; & \sum_{t=K+1}^{T} e_t e_{t-K} + \sum_{t=K+1}^{T} (\eta_t e_{t-K} + \eta_{t-K} e_t) + (\hat{\beta} - \beta)' \sum_{t=K+1}^{T} (X_t V_{t-K} + X_{t-K} V_t)(\hat{\Pi} - \Pi) \\ & - (\hat{\beta} - \beta)' \sum_{t=K+1}^{T} (X_t e_{t-K} + X_{t-K} e_t) - (\hat{\Pi} - \Pi)' \sum_{t=K+1}^{T} (V_t e_{t-K} + V_{t-K} e_t). \end{aligned} \tag{15}$$

The following theorem is obtained by applying Lemma 2 to the first term on the right-hand side of (12) whereas the remaining terms are shown to be negligible by directly applying the results of Saikkonen (1991), so that $\hat{C}_K \xrightarrow{d} N(0, \omega_a^2)$ under the null hypothesis. The consistency of $\hat{\omega}_a^2$ is also proved similarly to HML (2003). On the other hand, the test statistic diverges to infinity as proved by HML (2003) and then we omit the details.

Theorem 1 *Suppose that Assumptions 1, 2 and 3 hold. Under the null hypothesis, as* $T \to \infty$*,*

$$\hat{S}_K \to N(0, 1),$$

whereas under the fixed alternative, it diverges to infinity.

From Theorem 1, we can test for the null hypothesis of cointegration using the same test statistic as HML (2003) using the DOLS regression residuals, even though they are not causal but expressed as the leads and lags of the innovations.

3.3. Bias correction of the cointegration tests

As explained in the previous section, the cointegration test based on the autocovariance suffers from under-size distortion and we need to construct the bias-corrected version of the test statistic as suggest by HLM (2005). Because the first term on the right-hand side of (12) is the leading term, we define the bias term of (12) as the remaining terms up to $O_p(T^{-1/2})$ and the bias as its expectation up to $O(T^{-1/2})$. It is shown in the proof of Lemma 3 that the non-zero expectation from the bias term appears only from $R_{\beta,T}$ in (12) while $R_{\Pi,T}$ and R_T can be negligible.

Lemma 3 *The bias of* (12) *up to* $O(1/\sqrt{T})$*, denoted as* $-b$*, is given by*

$$-b = -\frac{(p_c + p_x)\omega_\eta^2}{\sqrt{T-K}\sigma_\eta^2}, \quad \textit{where } p_c = 1 \textit{ (constant case) or } p_c = 2 \textit{ (trend case).}$$

From the result of Lemma 3, the bias-corrected version of the test statistic is defined by

$$\tilde{S}_K = \frac{\tilde{C}_K + \tilde{b}}{\hat{\omega}_a} \quad \text{where} \quad \tilde{b} = \frac{(p_c + p_x)\hat{\omega}_\eta^2}{\sqrt{T-K}\hat{\sigma}_\eta^2}$$

with $\hat{\omega}_\eta^2$ is the long-run variance estimator based on $\hat{\eta}_t^*$ with the Bartlett kernel defined as (9) with $\tilde{a}_{K,t}$ replaced by $\hat{\eta}_t^*$. Then, we have the following corollary:

Corollary 1 *Suppose that Assumptions 1, 2 and 3 hold. Under the null hypothesis, as $T \to \infty$,*

$$\tilde{S}_K \to N(0,1),$$

whereas under the fixed alternative, it diverges to infinity.

3.4. Asymptotic behavior under the moderately integrated alternative

In this subsection, we investigate the effect of the choice of the lag order K on the power of the autocovariance based test. As seen in the proof of Lemma 2, the FCLT holds with the negligible terms proportional to $o(\sqrt{T}/K^2)$ and thus some appropriate divergence rate of K would be required to control the empirical size of the test. However, as proved below, the too fast divergence rate will result in an asymptotic loss of power.

To see the effect of the divergence rate of K on the asymptotic power, we extend model (6) as follows:

$$y_t = \beta' X_t + \sum_{j=-M}^{M} \pi_j' v_{t-j} + \eta_t^*, \qquad \eta_t^* = \dot{\eta}_t + e_t, \tag{16}$$

where $e_t = \sum_{j>|M|} \pi_j' v_{t-j}$ and

$$\dot{\eta}_t = \dot{\rho}\dot{\eta}_{t-1} + \xi_t \qquad \text{with} \qquad \dot{\rho} = 1 - \frac{c_1}{T^{\vartheta}} \qquad \text{for} \qquad 0 < \vartheta < 1, \tag{17}$$

c_1 is some positive constant and $\dot{\eta}_0 = 0$. In this model, we assume that $\{\xi_t\}$ is independent of $\{v_t\}$ and is an *i.i.d.* sequence with mean 0, variance σ_ξ^2, $E[\xi_t^3] = 0$ and finite fourth order moments. The condition of the third order moment is not necessarily required but the derivation becomes more complicated without this assumption. Note that $\dot{\eta}_t$ is different from the typical local to unity process, for which the autoregressive coefficient is defined as $1 - c_1/T$. In our case, $\dot{\rho}$ approaches 1 at a rate slower than the usual local alternative as T goes to infinity, because $0 < \vartheta < 1$. The autoregressive model as defined in (17) is sometimes called a moderately integrated or moderate deviation model and investigated by Giraitis and Phillips (2006) and Phillips and Magdalinos (2007a, b). The moderate deviation is also used for the investigation of cointegrated models by Kurozumi and Hayakawa (2009) and Magdalinos and Phillips (2009), and for moving average models by Yabe (2012). This device helps us understand the relation between K and the asymptotic power.

Theorem 2 *Suppose that Assumptions 1, 2 and 3 hold and that the bandwidth J is $o(T^{1/2})$ and J/K is* *bounded above. Then, for model* (16)*, as* $T \to \infty$*,*

$$T^{1/2(\vartheta-1)}\hat{S}_K = \begin{cases} o_p(1) & : \text{ if } \dfrac{K}{T^{\vartheta}} \to \infty \\ O_p(1) & : \text{ if } \dfrac{K}{T^{\vartheta}} \to \tau \quad (0 \le \tau < \infty), \end{cases}$$

where $O_p(1)$ *is in sharp order and the limit takes positive values.*

Theorem 2 implies that for a given value of ϑ, if the lag order K diverges to infinity at most as fast as the moderate deviation rate T^{ϑ}, then the autocovariance based test statistic diverges to infinity at a rate of $T^{(1-\vartheta)/2}$. However, if K goes to infinity faster than T^{ϑ}, then the test statistic does not diverge at this rate. This implies that if we choose large values of K relative to T^{θ}, then the test suffers from a loss of power asymptotically. For example, when $K_1 < K_2$, $K_1/K_2 \to 0$ and $K_1 = T^{\vartheta}$, $\hat{S}_{K_1}$ diverges to infinity at a rate of $T^{(1-\vartheta)/2}$ but $\hat{S}_{K_2}$ does not, so that $\hat{S}_{K_1}$ is asymptotically more powerful than $\hat{S}_{K_2}$.[8] In Section 5, we will observe this tendency in finite samples.

4. Panel Cointegration Test

In the case of panel cointegration, model (2) becomes

$$y_{i,t} = \beta_i' X_{i,t} + u_{i,t} \quad \text{for} \quad i = 1, 2, \cdots, N \quad \text{and} \quad t = 1, 2, \cdots, T, \tag{18}$$

where $X_{i,t} = [1, x_{i,t}']'$ (constant case) or $X_{i,t} = [1, t, x_{i,t}']'$ (trend case), $y_{i,t}$ and $x_{i,t}$ are 1- and $p_{i,x}$-dimensional processes with

$$x_{i,t} = x_{i,t-1} + v_{i,t} \quad \text{and} \quad u_{i,t} = \rho_i u_{i,t-1} + u_{i,t}^*.$$

Note that the specification of the non-stochastic term and the dimension of the $I(1)$ regressors can be different for individuals. Model (18) looks like just a simple extension from a univariate to a multivariate model, but this model is more complicated than the simple multivariate case in that we allow for cross-cointegration among regressors $x_{1,t}, \cdots, x_{N,t}$, which is likely the case because of the international/regional comovement.

Allowing for an abuse of notation, let $u_t^* = [u_{1,t}^*, u_{2,t}^*, \cdots, u_{N,t}^*]'$ and $v_t = [v_{1,t}', v_{2,t}', \cdots, v_{N,t}']'$ be N- and $p_x \equiv (p_{1,x} + p_{2,x} + \cdots + p_{N,x})$-dimensional vectors, respectively, which are defined in a different way from the previous section to save notation. In the case of panel cointegration, we make the following assumption:

[8] A similar result would be obtained for the case of the usual local alternative with $\vartheta = 1$ but the proof should be changed.

Assumption 1' *(a) $[u_t^{*\prime}, v_t']'$ is a vector linear process given by*

$$\left[\begin{array}{c} u_t^* \\ v_t \end{array} \right] = \sum_{j=0}^{\infty} \Phi_j \varepsilon_{t-j} \quad with \quad \sum_{j=0}^{\infty} j^2 \|\Phi_j\| < \infty, \tag{19}$$

where $\{\Phi_j\}$ is a set of $(N + p_x) \times p_\varepsilon$ coefficients (p_ε is not necessarily equal to $N + p_x$) and $\{\varepsilon_t\}$ is a p_ε-dimensional i.i.d. sequence with mean 0 and variance given by Σ_ε, which is positive definite, and has the finite fourth order moments.

(b) The marginal distribution of $[u_{i,t}^*, v_{i,t}']$ satisfies Assumption 1 for $i = 1, 2, \cdots, N$.

As in the univariate case, we do not allow for cointegration among regressors in each individual regression (18) by Assumption 1'(b). On the other hand, it is possible for some $x_{i,t}$ to be cointegrated with $x_{j,t}$ with $i \neq j$. In this case, because $x_{i,t}$ and $x_{j,t}$ are driven by common stochastic trends, the dimension of ε_t could be less than $N+p_x$ and hence Φ_j are not necessarily square matrices but the column dimension becomes smaller than the row dimension. We also note that the cross-sectional dependence in $u_{i,t}$ across i is allowed through the off-diagonal elements of Φ_j and Σ_ε and that common factors can be included in $u_{i,t}$ as far as they can be expressed as linear processes. Assumption 1'(b) implies that $[u_{i,t}^*, v_{i,t}']'$ can be expressed as in Assumption 1 using a $(p_{i,c} + p_{i,x})$-dimensional *i.i.d.* sequence $\{\varepsilon_{i,t}\}$ and that $\varepsilon_{i,s}$ are independent of $\varepsilon_{j,t}$ for $s \neq t$. The latter property will be used to establish the joint convergence of the individual test statistics.

As in Assumption 1', the error term has sometimes been supposed to be a linear process for the investigation of panel unit roots/cointegration with fixed N in the literature. For example, Chang (2004) assumed that the error term is an invertible linear process, which is basically the same as (19), and proposed to approximate it by an infinite order autoregressive process to use the sieve bootstrap method. Palm, Smeekes and Urbain (2011) considered the model with cross-sectional correlation generated from both the common factor and the variance of the error term. Since the factor model can be expressed as a linear process, their cross-sectional structure is also included in (19). As pointed out by Palm, Smeekes and Urbain (2011), the advantage of dealing with fixed N is that we can treat the flexible (cross) correlation structure, in both strong and weak form, although the distribution of the test statistic may depend on that structure in some cases, so that the bootstrap method have been used in such cases. However, as will be seen later, our test statistics are asymptotically free from the cross-sectional dependence and hence we can test for cointegration using the asymptotic critical values.

Let us see an example of the error structure. When $x_{i,t}$ is one-dimensional and common for all i and the errors are linear processes given by[9]

$$\Delta x_{i,t} = v_{i,t} = \sum_{j=0}^{\infty} \phi_j^v \varepsilon_{t-j}^v \quad \text{for all } i \quad \text{and} \quad u_{i,t}^* = \sum_{j=0}^{\infty} \phi_{i,j}^u \varepsilon_{i,t-j}^u,$$

where $\{\varepsilon_t^v\}$ is independent of $\{\varepsilon_{i,t}^u\}$, we can see that

$$\begin{bmatrix} u_{1,t}^* \\ \vdots \\ u_{N,t}^* \\ v_{1,t} \\ \vdots \\ v_{N,t} \end{bmatrix} = \sum_{j=0}^{\infty} \begin{bmatrix} \phi_{1,j}^u & 0 & \cdots & 0 & 0 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & 0 \\ 0 & \cdots & 0 & \phi_{N,j}^u & 0 \\ 0 & \cdots & 0 & 0 & \phi_j^v \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & \phi_j^v \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t-j}^u \\ \vdots \\ \varepsilon_{N,t-j}^u \\ \varepsilon_{t-j}^v \end{bmatrix} \quad \text{with} \quad \Sigma_{a,\varepsilon} = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1N} & \\ \vdots & \ddots & \vdots & 0 \\ \sigma_{N1} & \cdots & \sigma_{NN} & \\ & 0 & & \sigma_{vv} \end{bmatrix}.$$

In this case, $\{\varepsilon_t\}$ are $(N+1)$-dimensional vectors and $\{\Phi_j\}$ are $2N \times (N+1)$ matrices.

The null hypothesis in the panel case is that all the individuals are cointegrated whereas at least one individual is not cointegrated under the alternative. That is,

$$H_0: \ \rho_i < 1 \text{ for all } i \quad \text{vs.} \quad H_1: \ \rho_i = 1 \text{ for } i = 1, \cdots, N_1 \text{ with } 1 \le N_1 \le N.$$

Note that because the cross-sectional dimension N is fixed in our model, we can reject the null hypothesis even if only one individual is not cointegrated. However, it is not difficult to expect that the test against small N_1 would be less powerful than that against large N_1.

As in the univariate case, individual regression (18) is augmented by the leads and lags of the first differences of the $I(1)$ regressors and we obtain the DOLS regression given by

$$y_{i,t} = \beta_i' X_{i,t} + \sum_{j=-M}^{M} \pi_{i,j}' v_{i,t-j} + \eta_{i,t}^* \tag{20}$$

where $\eta_{i,t}^*$ is defined as before and the standardized regression residuals are defined as

$$\tilde{\eta}_{i,t}^* = \frac{\hat{\eta}_{i,t}^*}{\hat{\sigma}_{i,\eta}} \quad \text{where} \quad \hat{\sigma}_{i,\eta}^2 = \frac{1}{T} \sum_{t=1}^{T} \hat{\eta}_{i,t}^{*2}.$$

Note that the truncation point M can be different over cross-sections but we proceed with the same M for notational convenience.

In this case, the test statistic for panel cointegration is given by

$$\hat{S}_K = \frac{\tilde{C}_K}{\hat{\omega}_a} \quad \text{where} \quad \tilde{C}_K = \frac{1}{\sqrt{T-K}} \sum_{t=K+1}^{T} \tilde{a}_{K,t} \quad \text{with} \quad \tilde{a}_{K,t} = \sum_{i=1}^{N} \tilde{\eta}_{i,t}^* \tilde{\eta}_{i,t-K}^*,$$

[9] $u_{i,t}^*$ could be expressed in more general way such as $u_t^* = \sum_{j=0}^{\infty} \Phi_j^u \varepsilon_{t-j}^u$ with $\{\Phi_j^u\}$ are not necessarily diagonal. We give a simple example with diagonal Φ_j to illustrate how the dimension of ε_t could be smaller than $N + p_x$.

while the bias-corrected version of the test statistic is defined as

$$\tilde{S}_K = \frac{\tilde{C}_K + \tilde{b}}{\hat{\omega}_a} \quad \text{where} \quad \tilde{b} = \frac{1}{\sqrt{T-K}} \sum_{i=1}^{N} \frac{(p_{i,c} + p_{i,x})\hat{\omega}_{i,\eta}^2}{\hat{\sigma}_{i,\eta}^2},$$

where $\hat{\omega}_a^2$ is defined by (9) using $\tilde{a}_{K,t}$, the cross-sectional sum of individual autocovariances.

Theorem 3 *Suppose that Assumptions 1', 2 and 3 hold. Under the null hypothesis, as $T \to \infty$,*

$$\hat{S}_K,\ \tilde{S}_K \to N(0,1),$$

whereas under the fixed alternative, they diverge to infinity.

Remark 1 *The test statistics $\hat{S}_K$ and $\tilde{S}_K$ are based on a simple average of the autocovariances of individuals, but it may be possible to test for panel cointegration based on the weighted sum of the autocovariances, which might lead to the improvement of the power of the tests. However, it seems there is no guidance for how to choose the weights. Since the regression errors are normalized by the standard deviation, it is natural to choose the uniform weights as used for our tests. We may also consider a test statistic based on the maximal autocovariance, as pointed out by a referee. However, because the autocovariances are correlated across cross-sections, the limiting distribution of such a test statistic does depend on the correlation structure. The bootstrap method may be one of the possible solutions but it is beyond our scope of this paper.*

Remark 2 *Theorem 3 holds when the individuals are independent, in which case the test statistics could be much simpler and it would be possible to modify the test statistics to accommodate the large N and large T asymptotics (the joint asymptotics). This implies that if the cross-correlation structure can be estimated and removed, then we could consider models with the cross-sectional dimension N being as large as or larger than T. This would be an interesting extension but we restrict our attention to the current case with the fixed N and large T asymptotics to allow for a wide class of the dependence structure.*

As discussed in the introduction, the advantage of using HLM (2005) test is that we do not have to rely on the joint limit theorem in order to obtain a test statistic whose null limiting distribution is free of nuisance parameter. This is because the test statistic in the univariate case has the limiting normal distribution. As a result, we can apply our test even for panel data with small N.

5. Monte Carlo Simulations

5.1. Single cointegration tests

In this section, we first investigate the finite sample performance of the single cointegration tests proposed in this paper. The data generating process is given by

$$y_t = \alpha' c_t + \beta x_t + u_t, \qquad x_t = x_{t-1} + v_t,$$

$$u_t = \phi u_{t-1} + \varepsilon_t^u, \qquad v_t = \psi v_{t-1} + \varepsilon_t^v,$$

where $c_t = 1$ in the constant case while $c_t = [1, t]'$ in the trend case, $\alpha = 0$, $\beta = 1$, and $[\varepsilon_t^u, \varepsilon_t^v]' \sim i.i.d.N(0, \Sigma)$ with $vech(\Sigma) = [1, 0.5, 1]'$. The initial values of u_t and v_t are $u_0 = v_0 = 0$. To control serial correlation in v_t, we set $\psi = 0$, 0.4, 0.8. Under the null hypothesis of cointegration, ϕ must be less than 1 in absolute value and then we consider three cases; $\phi = 0$, 0.4, 0.8, while the alternative of no cointegration corresponds to the case with $\phi = 1$.

Throughout the simulations, the bandwidth J for the long-run variance estimation and the leads-lags truncation parameter M are set to[10]

$$J = \left[12(T/100)^{1/4}\right] \quad \text{and} \quad M = \left[2(T/100)^{1/5}\right].$$

We also investigate the effect of the lag order K on our autocovariance based tests because the finite sample performance will crucially depend on K. We consider

$$K = \left[(aT)^{\delta}\right] \quad \text{for } a = 1,\ 2 \text{ and } 3 \text{ and } \delta = 1/4,\ 1/2 \text{ and } 3/4,$$

so that 9 lag orders are used in the simulations. Note that HML (2003) and HLM (2005) recommended $K = O(T^{1/2})$. In addition to the autocovariance based test ($\hat{S}_K$) and its bias corrected version ($\tilde{S}_K$) using the DOLS residuals, we calculate the autocovariance based test using the OLS residuals ($\hat{S}_K^{ols}$). As discussed in Section 3.1, $\hat{S}_K^{ols}$ is asymptotically normally distributed under the null hypothesis and thus it should be compared with the DOLS-based version. The number of replications is 5,000 and the significance level is set to 0.05. All computations are conducted using the GAUSS matrix language.

For the purpose of comparison, we also calculate the single cointegration test statistic proposed by Shin (1994), which is one of the most frequently used test in applications, and the LBIU test by Kurozumi and Arai (2008), which can control the empirical size well even when the errors are strongly serially correlated. The leads-lags truncation parameter for the Shin's test is the same as the above, while the semiparametric correction is used for the LBIU test; see Kurozumi and Arai (2008) for details.

[10] We also conducted simulations with $M = [4(T/100)^{1/5}]$, $M = [2(T/100)^{1/6}]$, $M = [4(T/100)^{1/6}]$. As a whole, the longer leads and lags result in the empirical size slightly closer to the nominal one but the power is reduced. We chose $M = [2(T/100)^{1/5}]$ because the empirical size is sufficiently close to the nominal one compared to the other choices while the test with this choice has sufficient power.

Table 1 reports the rejection frequencies of the tests. The Shin's test can well control the size of the test when the serial correlation is not strong but when $\phi = 0.8$, it suffers from over-size distortion. On the other hand, the empirical size of the LBIU test is close to 0.05 even in the case of strong serial correlation. For the autocovariance based tests, the columns $\hat{S}_K^{ols}(a)$, $\hat{S}_K(a)$ and $\tilde{S}_K(a)$ correspond to the case when $K = [(aT)^{1/2}]$ for $a = 1$, 2 and 3 are used; the cases with the other lag orders are omitted to save space.[11] From the table, we observe that the tests with no bias correction, both $\hat{S}_K^{ols}(a)$ and $\hat{S}_K(a)$, tend to be conservative because of the effect of the negative bias, although the OLS based test performs slightly better than the DOLS based test. On the other hand, the empirical sizes of the bias-corrected versions, $\tilde{S}_K(a)$, are close to the nominal one except for the case where $a = 1$ and $\phi = \psi = 0.8$. Overall, the finite sample performance of our test under the null hypothesis is better than that of the Shin's test and as good as the LBIU test.

With respect to power, the Shin's test seems more powerful than the LBIU test by observing the results for $\psi = 0$ and 0.4, in which case the sizes of these two tests are close to the nominal one. On the other hand, the power of the bias-corrected version of the autocovariance based test depends on the choice of K; the test tends to be more powerful than the Shin's test for smaller K whereas it is less powerful in the most generous case. Among the autocovariance based tests, $\tilde{S}_K(a)$ is most powerful than the other two tests and the OLS based test is the second best for a given lag order K, although the difference between $\hat{S}_K^{ols}(a)$ and $\hat{S}_K(a)$ is slight.

As a whole, the simulation results suggest that we should carefully choose the lag order K. Taking into account the finite sample performance under both the null and the alternative hypotheses, we recommend using our bias-corrected test with $K = [(2T)^{1/2}]$ or $K = [(3T)^{1/2}]$, although the power drops dramatically from $a = 1$ to 2.

5.2. Panel cointegration tests

We next investigate the finite sample performance of the panel cointegration tests. The data generating process is similar to the single equation case and it is given by

$$y_{i,t} = \alpha_i' c_t + \beta_i x_{i,t} + e_{i,t}, \qquad x_{i,t} = x_{i,t} + v_{i,t}, \qquad e_{i,t} = u_{i,t} + \lambda_i f_t,$$

$$u_{i,t} = \phi_i u_{i,t-1} + \varepsilon_{i,t}^u, \qquad v_{i,t} = \psi_i v_{i,t-1} + \varepsilon_{i,t}^v,$$

[11] Roughly speaking, the autocovariance based tests with $K = [(aT)^{1/4}]$ and $[(aT)^{3/4}]$ result in the over-size distortion when the errors are strongly serially correlated.

Table 1: Empirical size and power of univariate tests

T	ψ	ϕ	Shin	LBIU	$\hat{S}_K^{ols}(1)$	$\hat{S}_K^{ols}(2)$	$\hat{S}_K^{ols}(3)$	$\hat{S}_K(1)$	$\hat{S}_K(2)$	$\hat{S}_K(3)$	$\tilde{S}_K(1)$	$\tilde{S}_K(2)$	$\tilde{S}_K(3)$
			costant model										
100			0.052	0.061	0.029	0.030	0.034	0.024	0.021	0.027	0.045	0.040	0.048
300	0.0	0.0	0.048	0.056	0.044	0.038	0.040	0.039	0.042	0.036	0.050	0.052	0.047
500			0.047	0.053	0.042	0.042	0.044	0.038	0.037	0.040	0.048	0.047	0.049
100			0.397	0.157	0.414	0.200	0.129	0.271	0.113	0.069	0.712	0.477	0.344
300	0.0	1.0	0.720	0.518	0.870	0.676	0.537	0.853	0.642	0.504	0.928	0.783	0.669
500			0.840	0.721	0.971	0.870	0.767	0.967	0.858	0.751	0.982	0.917	0.833
100			0.061	0.056	0.021	0.020	0.031	0.017	0.016	0.020	0.058	0.049	0.057
300	0.4	0.4	0.055	0.054	0.038	0.035	0.033	0.033	0.034	0.033	0.055	0.054	0.054
500			0.055	0.052	0.032	0.034	0.039	0.030	0.034	0.036	0.048	0.049	0.053
100			0.386	0.133	0.406	0.192	0.122	0.257	0.104	0.061	0.698	0.456	0.330
300	0.4	1.0	0.715	0.473	0.867	0.669	0.533	0.852	0.636	0.497	0.928	0.781	0.666
500			0.838	0.688	0.970	0.868	0.766	0.967	0.858	0.749	0.982	0.916	0.831
100			0.082	0.066	0.029	0.016	0.015	0.014	0.008	0.008	0.141	0.075	0.065
300	0.8	0.8	0.088	0.055	0.034	0.027	0.025	0.028	0.022	0.023	0.070	0.059	0.058
500			0.089	0.058	0.031	0.031	0.033	0.023	0.026	0.029	0.055	0.053	0.057
100			0.334	0.082	0.380	0.165	0.104	0.209	0.074	0.044	0.641	0.393	0.273
300	0.8	1.0	0.697	0.321	0.860	0.652	0.515	0.842	0.614	0.469	0.926	0.763	0.645
500			0.829	0.544	0.967	0.866	0.755	0.966	0.851	0.737	0.984	0.912	0.824
			trend model										
T	ψ	ϕ	Shin	LBIU	$\hat{S}_K^{ols}(1)$	$\hat{S}_K^{ols}(2)$	$\hat{S}_K^{ols}(3)$	$\hat{S}_K(1)$	$\hat{S}_K(2)$	$\hat{S}_K(3)$	$\tilde{S}_K(1)$	$\tilde{S}_K(2)$	$\tilde{S}_K(3)$
100			0.063	0.078	0.025	0.025	0.029	0.021	0.020	0.024	0.047	0.045	0.052
300	0.0	0.0	0.050	0.052	0.041	0.036	0.035	0.036	0.039	0.033	0.052	0.054	0.051
500			0.049	0.052	0.039	0.040	0.041	0.035	0.034	0.039	0.049	0.049	0.050
100			0.369	0.054	0.156	0.017	0.002	0.076	0.004	0.001	0.532	0.256	0.145
300	0.0	1.0	0.762	0.311	0.677	0.377	0.231	0.636	0.339	0.202	0.847	0.597	0.434
500			0.879	0.557	0.892	0.665	0.493	0.879	0.644	0.467	0.949	0.794	0.637
100			0.072	0.069	0.016	0.017	0.024	0.012	0.014	0.019	0.059	0.050	0.060
300	0.4	0.4	0.058	0.051	0.029	0.030	0.027	0.026	0.029	0.028	0.055	0.057	0.056
500			0.057	0.053	0.027	0.029	0.034	0.027	0.029	0.031	0.048	0.052	0.056
100			0.356	0.049	0.147	0.015	0.003	0.069	0.004	0.002	0.513	0.239	0.141
300	0.4	1.0	0.759	0.264	0.670	0.370	0.227	0.632	0.331	0.197	0.843	0.590	0.429
500			0.876	0.515	0.889	0.662	0.490	0.878	0.641	0.463	0.949	0.789	0.632
100			0.094	0.069	0.012	0.005	0.008	0.005	0.003	0.007	0.116	0.063	0.060
300	0.8	0.8	0.107	0.050	0.019	0.017	0.016	0.016	0.015	0.016	0.067	0.059	0.064
500			0.101	0.051	0.021	0.020	0.025	0.016	0.016	0.020	0.052	0.054	0.060
100			0.284	0.053	0.129	0.011	0.001	0.050	0.003	0.003	0.435	0.184	0.109
300	0.8	1.0	0.738	0.138	0.650	0.342	0.210	0.601	0.301	0.173	0.821	0.557	0.400
500			0.868	0.355	0.877	0.647	0.471	0.871	0.623	0.441	0.945	0.777	0.618

where $\alpha_i = 0$, $\beta_i = 1$, and $[\varepsilon^u_{i,t}, \varepsilon^v_{i,t}]' \sim i.i.d.N(0, \Sigma)$ with $vech(\Sigma) = [1, 0.5, 1]'$ with $u_0 = v_0 = 0$. The error terms $e_{i,t}$ consist of the idiosyncratic errors $u_{i,t}$ and the common components $\lambda_i f_t$ with common factor f_t and loading factors λ_i. The idiosyncratic errors, $u_{i,t}$, and the driving force of the I(1) regressors, $v_{i,t}$, are correlated for the same individual i, but they are cross-sectionally independent. We set $\lambda_i = 0$ for the case of no cross-sectional correlation while $\lambda_i \sim U(0, 1)$ and $f_t \sim i.i.d.N(0, 1)$ for the case of cross-sectional dependence. To see the effect of serial correlation on the tests under the null hypothesis, we consider three cases; $\phi_i \sim U(-0.4, 0.4)$ and $\psi_i \sim U(-0.4, 0.4)$ (mild serial correlation), $\phi_i \sim U(-0.8, 0.8)$ and $\psi_i \sim U(-0.8, 0.8)$ (diversified serial correlation), and $\phi_i \sim U(0.7, 0.9)$ and $\psi_i \sim U(0.7, 0.9)$ (strong serial correlation). Under the alternative hypothesis that not all the individuals are cointegrated, we generate $\phi_i = 1$ for $i = 1, \cdots, N_1$ and ϕ_i for $i = N_1 + 1, \cdots, N$ are the same as in the null hypothesis.

Tables 2–4 correspond to the case of cross-sectional dependence. Table 2 summarizes the results for the case of mild serial correlation. For the purpose of comparison, we also calculate the Lagrange multiplier (LM) test proposed by McCoskey and Kao (1998) and report the rejection frequencies on the column denoted by LM.[12] Since the original LM test assumes the cross-sectional dependence, we did not use the asymptotic critical values for the LM test but obtained the critical values by the sieve bootstrap method as suggested by Westerlund and Edgerton (2007). From the rows of $N_1/N = 0$ in the table, we can see that the LM test controls the empirical size very well in almost all the cases. As theoretically expected from the previous section, the autocovariance based tests with no bias correction tend to under-reject the null hypothesis. In particular, the empirical size of the test based on the DOLS residuals is almost less than 0.01 when $N = 100$. On the other hand, the bias-corrected versions work well, which indicates the effectiveness of our bias correction, except for the case where $T = 100$ and $N \geq 50$ in the trend case. With respect to power, the bootstrap LM test is much less powerful than our test with bias correction in many cases, except for the trend case with $T = 100$ and $N_1/N = 0.5$. In general, the power of the bootstrapped version of the LM test increases relatively slowly as T increases, whereas the powers of $\tilde{S}_k(a)$ is generally good; it increases with the sample size T and the ratio N_1/N as expected.

The relative performance in the diversified serial correlation case in Table 3 is essentially similar to the mild serial correlation case in Table 2, but the strong serial correlation case in Table 4 is different from the former two cases. The empirical sizes of the test with no bias correction is zero in many cases while $\tilde{S}_k(1)$ suffers from over-size distortion. This is surprising because $\tilde{S}_k(1)$ performs very well in the

[12] We used the asymptotic mean and variance to construct the test statistic by McCoskey and Kao (1998). Because they reported the mean and variance only in the constant case, we calculate those values in the trend case by simulations.

time series case as in Table 1, from which we could expect that it should also work well in panel data. It is difficult to find a theoretical reason, but it seems that the long-run variance estimator might badly affect the finite sample performance in the panel case. On the other hand, $\tilde{S}_k(2)$ and $\tilde{S}_k(3)$ work well under both the null and the alternative in the constant case while they tend to be under-reject in many cases in the trend case. With respect to the LM test, the empirical size is close to the nominal one while it is not as powerful as $\tilde{S}_K(a)$ in many cases.

In the case of independent cross-sections, we obtained similar results, which are omitted to save space (details are available upon request). It seems that the cross-sectional dependence has only a minor effect on our test statistics.

The above results correspond to the case where $x_{i,t}$ is cross-sectionally independent. However, in practical analysis, they may be correlated and moreover, it is possible for $x_{i,t}$ to be cointegrated with $x_{j,t}$. Then, we consider the same data generating process as before except that

$$x_{i,t} = \gamma_i(x_t + w_{i,t}), \qquad x_t = x_{t-1} + v_t,$$

where $v_t = \psi v_{t-1} + \varepsilon_t^v$, $w_{i,t} \sim i.i.d.N(0,1)$ and $\gamma_i \sim U(0.5, 1.5)$. That is, we consider the case where $x_{i,t}$ are driven by the common stochastic trend, which implies that $[x_{1,t}, \cdots, x_{N,t}]'$ are cointegrated with cointegrating rank $N-1$.

Table 5 corresponds to the cross-cointegration case with cross-sectional correlation and diversified serial correlation. In this case, the bias-corrected tests tend to over-reject the null hypothesis compared to the case with no cross-cointegration. In particular, $\tilde{S}_k(3)$ suffers from size distortion even $T = 300$ and large N. On the other hand, $\tilde{S}_k(1)$ and $\tilde{S}_k(2)$ performs well except for the case when $T = 100$ and N is not small. The relative performance of the LM test to $\tilde{S}_K(a)$ is similar to Tables 2–4.

When the cross-cointegration is allowed with the strong serial correlation, $\tilde{S}_k(1)$ severely suffers from over-size distortion; the empirical size is no less than 0.11 and it reaches more than 0.5 in some cases. The empirical size of $\tilde{S}_k(2)$ also tends to be more than 0.15 when $T = 100$ but it gets close to the nominal one for $T = 300$ and $T = 500$. As a whole, the size distortion of the bias-corrected tests is milder in the trend case than in the constant case. Details are available upon request.

In summary, our bias-corrected tests with $K = [(2T)^{1/2}]$ or $K = [(3T)^{1/2}]$ are recommended in practical analysis on the basis of our extensive simulations, although the power decreases from $a = 1$ to 2 as in the univariate case.

Table 2: Empirical size and power of panel tests
(cross correlation, mild serial correlation)

			costant model									
T	N	N_1/N	LM	$\hat{S}_K^{ols}(1)$	$\hat{S}_K^{ols}(2)$	$\hat{S}_K^{ols}(3)$	$\hat{S}_K(1)$	$\hat{S}_K(2)$	$\hat{S}_K(3)$	$\tilde{S}_K(1)$	$\tilde{S}_K(2)$	$\tilde{S}_K(3)$
100	10		0.045	0.015	0.027	0.023	0.012	0.015	0.019	0.060	0.067	0.074
100	25	0	0.041	0.013	0.014	0.020	0.006	0.011	0.013	0.058	0.074	0.089
100	50		0.047	0.014	0.014	0.017	0.005	0.008	0.010	0.067	0.088	0.094
100	100		0.072	0.016	0.010	0.019	0.005	0.011	0.009	0.064	0.086	0.091
300	10		0.062	0.033	0.036	0.020	0.026	0.030	0.027	0.053	0.055	0.056
300	25	0	0.054	0.027	0.026	0.025	0.022	0.024	0.021	0.056	0.061	0.062
300	50		0.044	0.028	0.021	0.024	0.019	0.017	0.018	0.060	0.053	0.055
300	100		0.050	0.023	0.023	0.025	0.015	0.018	0.023	0.055	0.057	0.066
500	10		0.053	0.031	0.030	0.051	0.033	0.032	0.036	0.054	0.051	0.055
500	25	0	0.051	0.033	0.034	0.035	0.025	0.025	0.029	0.051	0.055	0.056
500	50		0.056	0.025	0.024	0.026	0.021	0.026	0.028	0.051	0.060	0.062
500	100		0.055	0.027	0.030	0.034	0.026	0.025	0.023	0.057	0.055	0.058
100	10		0.160	0.305	0.138	0.071	0.175	0.070	0.038	0.622	0.427	0.311
100	25	0.2	0.174	0.453	0.163	0.083	0.244	0.072	0.026	0.816	0.598	0.433
100	50		0.271	0.582	0.181	0.067	0.307	0.074	0.020	0.921	0.703	0.513
100	100		0.337	0.729	0.243	0.079	0.397	0.084	0.018	0.960	0.780	0.585
300	10		0.236	0.899	0.707	0.541	0.865	0.646	0.482	0.963	0.847	0.737
300	25	0.2	0.375	0.999	0.921	0.784	0.989	0.877	0.715	0.999	0.983	0.934
300	50		0.513	1.000	0.987	0.918	0.999	0.958	0.835	1.000	0.999	0.985
300	100		0.678	1.000	1.000	0.982	1.000	0.993	0.924	1.000	1.000	0.997
500	10		0.284	0.995	0.909	0.814	0.985	0.911	0.807	0.996	0.965	0.911
500	25	0.2	0.447	1.000	1.000	0.983	1.000	0.995	0.968	1.000	1.000	0.995
500	50		0.633	1.000	1.000	0.999	1.000	1.000	0.997	1.000	1.000	1.000
500	100		0.794	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
100	10		0.345	0.857	0.404	0.146	0.665	0.212	0.067	0.980	0.814	0.612
100	25	0.5	0.531	0.978	0.578	0.194	0.876	0.277	0.061	1.000	0.960	0.802
100	50		0.770	1.000	0.811	0.263	0.984	0.394	0.052	1.000	0.997	0.934
100	100		0.929	1.000	0.920	0.366	0.999	0.517	0.050	1.000	1.000	0.984
300	10		0.468	1.000	0.985	0.926	1.000	0.977	0.886	1.000	0.999	0.986
300	25	0.5	0.769	1.000	1.000	0.996	1.000	1.000	0.991	1.000	1.000	1.000
300	50		0.940	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
300	100		0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
500	10		0.531	1.000	1.000	0.996	1.000	1.000	0.995	1.000	1.000	0.999
500	25	0.5	0.798	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
500	50		0.960	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
500	100		0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 2: (continued)
(cross correlation, mild serial correlation)

			trend model									
T	N	N_1/N	LM	$\hat{S}_K^{ols}(1)$	$\hat{S}_K^{ols}(2)$	$\hat{S}_K^{ols}(3)$	$\hat{S}_K(1)$	$\hat{S}_K(2)$	$\hat{S}_K(3)$	$\tilde{S}_K(1)$	$\tilde{S}_K(2)$	$\tilde{S}_K(3)$
100	10		0.065	0.014	0.014	0.021	0.007	0.012	0.018	0.063	0.076	0.089
100	25	0	0.049	0.006	0.005	0.006	0.004	0.009	0.011	0.060	0.088	0.114
100	50		0.062	0.003	0.004	0.006	0.002	0.005	0.007	0.068	0.106	0.122
100	100		0.055	0.007	0.006	0.010	0.003	0.007	0.008	0.065	0.100	0.119
300	10		0.058	0.027	0.025	0.019	0.022	0.027	0.025	0.055	0.061	0.064
300	25	0	0.058	0.032	0.016	0.025	0.015	0.020	0.017	0.061	0.071	0.072
300	50		0.045	0.016	0.013	0.013	0.015	0.012	0.016	0.065	0.063	0.068
300	100		0.050	0.015	0.016	0.014	0.011	0.013	0.016	0.058	0.068	0.079
500	10		0.071	0.024	0.029	0.033	0.028	0.027	0.031	0.059	0.056	0.058
500	25	0	0.060	0.037	0.029	0.022	0.021	0.021	0.022	0.055	0.058	0.063
500	50		0.056	0.021	0.023	0.026	0.016	0.019	0.022	0.055	0.065	0.069
500	100		0.055	0.022	0.020	0.028	0.019	0.019	0.018	0.061	0.063	0.067
100	10		0.115	0.064	0.008	0.004	0.028	0.009	0.003	0.399	0.211	0.152
100	25	0.2	0.149	0.063	0.007	0.000	0.024	0.003	0.002	0.522	0.268	0.180
100	50		0.157	0.057	0.003	0.001	0.017	0.001	0.000	0.610	0.308	0.192
100	100		0.181	0.076	0.002	0.001	0.018	0.001	0.000	0.690	0.333	0.196
300	10		0.215	0.643	0.321	0.154	0.573	0.257	0.123	0.875	0.635	0.454
300	25	0.2	0.314	0.898	0.454	0.165	0.817	0.359	0.148	0.987	0.865	0.661
300	50		0.371	0.975	0.604	0.215	0.932	0.425	0.136	0.999	0.953	0.783
300	100		0.419	1.000	0.729	0.252	0.978	0.534	0.166	1.000	0.982	0.870
500	10		0.250	0.953	0.707	0.487	0.905	0.652	0.443	0.976	0.859	0.713
500	25	0.2	0.383	0.998	0.933	0.705	0.995	0.874	0.622	1.000	0.986	0.920
500	50		0.528	1.000	0.987	0.854	1.000	0.971	0.761	1.000	0.999	0.985
500	100		0.655	1.000	0.999	0.951	1.000	0.995	0.882	1.000	1.000	0.998
100	10		0.259	0.252	0.005	0.000	0.097	0.003	0.000	0.821	0.390	0.183
100	25	0.5	0.385	0.318	0.002	0.000	0.081	0.000	0.000	0.940	0.482	0.198
100	50		0.550	0.402	0.000	0.000	0.078	0.000	0.000	0.993	0.565	0.193
100	100		0.701	0.517	0.000	0.000	0.099	0.000	0.000	0.999	0.647	0.187
300	10		0.483	0.994	0.711	0.340	0.982	0.628	0.264	0.999	0.953	0.802
300	25	0.5	0.727	1.000	0.896	0.485	1.000	0.851	0.369	1.000	0.999	0.961
300	50		0.922	1.000	0.989	0.649	1.000	0.970	0.464	1.000	1.000	0.999
300	100		0.984	1.000	0.999	0.830	1.000	0.998	0.600	1.000	1.000	1.000
500	10		0.551	1.000	0.980	0.856	1.000	0.969	0.806	1.000	0.997	0.966
500	25	0.5	0.817	1.000	1.000	0.978	1.000	1.000	0.969	1.000	1.000	0.999
500	50		0.961	1.000	1.000	1.000	1.000	1.000	0.999	1.000	1.000	1.000
500	100		1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3: Empirical size and power of panel tests
(cross correlation, diversified serial correlation)

			constant model									
T	N	N_1/N	LM	$\hat{S}_K^{ols}(1)$	$\hat{S}_K^{ols}(2)$	$\hat{S}_K^{ols}(3)$	$\hat{S}_K(1)$	$\hat{S}_K(2)$	$\hat{S}_K(3)$	$\tilde{S}_K(1)$	$\tilde{S}_K(2)$	$\tilde{S}_K(3)$
100	10		0.053	0.024	0.023	0.029	0.012	0.014	0.016	0.059	0.068	0.077
100	25	0	0.040	0.008	0.008	0.013	0.004	0.007	0.008	0.069	0.085	0.098
100	50		0.044	0.005	0.004	0.011	0.002	0.003	0.005	0.072	0.092	0.110
100	100		0.068	0.007	0.004	0.013	0.002	0.005	0.006	0.046	0.078	0.105
300	10		0.061	0.030	0.038	0.025	0.028	0.029	0.028	0.059	0.058	0.059
300	25	0	0.056	0.028	0.018	0.016	0.015	0.017	0.016	0.059	0.066	0.065
300	50		0.039	0.015	0.011	0.015	0.011	0.009	0.011	0.066	0.059	0.061
300	100		0.060	0.016	0.019	0.021	0.009	0.013	0.017	0.052	0.060	0.070
500	10		0.053	0.032	0.034	0.049	0.032	0.033	0.035	0.058	0.057	0.056
500	25	0	0.050	0.040	0.031	0.030	0.021	0.021	0.022	0.056	0.059	0.060
500	50		0.057	0.028	0.018	0.026	0.013	0.015	0.018	0.058	0.058	0.067
500	100		0.056	0.022	0.025	0.026	0.021	0.019	0.018	0.059	0.058	0.064
100	10		0.160	0.293	0.115	0.060	0.151	0.056	0.030	0.605	0.395	0.301
100	25	0.2	0.178	0.405	0.124	0.052	0.192	0.049	0.019	0.812	0.570	0.427
100	50		0.272	0.522	0.129	0.039	0.229	0.040	0.010	0.920	0.679	0.505
100	100		0.343	0.690	0.212	0.072	0.333	0.060	0.014	0.961	0.781	0.598
300	10		0.239	0.868	0.672	0.505	0.836	0.602	0.442	0.956	0.832	0.716
300	25	0.2	0.382	0.998	0.897	0.740	0.983	0.844	0.664	0.999	0.980	0.929
300	50		0.494	1.000	0.982	0.882	0.999	0.946	0.789	1.000	0.998	0.987
300	100		0.701	1.000	1.000	0.983	1.000	0.995	0.928	1.000	1.000	0.999
500	10		0.272	0.989	0.886	0.776	0.980	0.889	0.787	0.995	0.959	0.896
500	25	0.2	0.443	1.000	1.000	0.983	1.000	0.995	0.960	1.000	1.000	0.994
500	50		0.645	1.000	1.000	0.999	1.000	1.000	0.997	1.000	1.000	1.000
500	100		0.800	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
100	10		0.341	0.856	0.399	0.163	0.646	0.207	0.068	0.973	0.805	0.602
100	25	0.5	0.539	0.971	0.522	0.188	0.834	0.218	0.049	0.999	0.945	0.785
100	50		0.768	1.000	0.742	0.221	0.973	0.313	0.035	1.000	0.996	0.928
100	100		0.931	1.000	0.914	0.346	0.997	0.463	0.037	1.000	1.000	0.983
300	10		0.471	1.000	0.986	0.924	1.000	0.976	0.884	1.000	0.999	0.984
300	25	0.5	0.762	1.000	1.000	0.996	1.000	1.000	0.987	1.000	1.000	1.000
300	50		0.933	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
300	100		0.992	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
500	10		0.525	1.000	1.000	0.995	1.000	1.000	0.994	1.000	1.000	0.999
500	25	0.5	0.802	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
500	50		0.961	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
500	100		0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3: (continued)
(cross correlation, diversified serial correlation)

			trend model									
T	N	N_1/N	LM	$\hat{S}_K^{ols}(1)$	$\hat{S}_K^{ols}(2)$	$\hat{S}_K^{ols}(3)$	$\hat{S}_K(1)$	$\hat{S}_K(2)$	$\hat{S}_K(3)$	$\tilde{S}_K(1)$	$\tilde{S}_K(2)$	$\tilde{S}_K(3)$
100	10		0.059	0.011	0.015	0.019	0.008	0.010	0.013	0.057	0.077	0.091
100	25	0	0.061	0.001	0.002	0.000	0.002	0.005	0.005	0.060	0.089	0.122
100	50		0.046	0.000	0.000	0.002	0.000	0.001	0.004	0.056	0.093	0.132
100	100		0.070	0.002	0.000	0.003	0.001	0.002	0.005	0.041	0.090	0.136
300	10		0.058	0.024	0.030	0.027	0.021	0.024	0.021	0.058	0.063	0.066
300	25	0	0.060	0.018	0.012	0.015	0.009	0.011	0.010	0.059	0.076	0.078
300	50		0.060	0.008	0.002	0.008	0.003	0.004	0.007	0.066	0.067	0.076
300	100		0.066	0.012	0.010	0.007	0.004	0.008	0.010	0.053	0.072	0.084
500	10		0.075	0.016	0.027	0.036	0.027	0.028	0.030	0.060	0.059	0.062
500	25	0	0.059	0.036	0.029	0.030	0.013	0.013	0.015	0.061	0.060	0.067
500	50		0.059	0.014	0.018	0.019	0.007	0.009	0.010	0.061	0.065	0.075
500	100		0.059	0.013	0.012	0.014	0.014	0.012	0.012	0.062	0.064	0.071
100	10		0.123	0.048	0.008	0.007	0.020	0.005	0.003	0.371	0.198	0.142
100	25	0.2	0.156	0.033	0.002	0.000	0.010	0.001	0.000	0.490	0.241	0.171
100	50		0.146	0.033	0.000	0.001	0.004	0.000	0.000	0.548	0.254	0.170
100	100		0.224	0.044	0.001	0.001	0.007	0.001	0.000	0.639	0.301	0.189
300	10		0.202	0.597	0.296	0.135	0.522	0.226	0.107	0.858	0.603	0.432
300	25	0.2	0.318	0.849	0.391	0.130	0.755	0.294	0.108	0.986	0.851	0.646
300	50		0.400	0.963	0.487	0.137	0.899	0.331	0.079	1.000	0.952	0.787
300	100		0.458	1.000	0.707	0.224	0.977	0.513	0.144	1.000	0.987	0.886
500	10		0.240	0.919	0.636	0.444	0.882	0.609	0.402	0.972	0.840	0.696
500	25	0.2	0.387	0.998	0.925	0.694	0.992	0.841	0.567	1.000	0.984	0.913
500	50		0.549	1.000	0.982	0.831	1.000	0.956	0.711	1.000	0.999	0.985
500	100		0.661	1.000	0.999	0.947	1.000	0.996	0.887	1.000	1.000	0.999
100	10		0.261	0.235	0.008	0.001	0.089	0.004	0.000	0.805	0.378	0.181
100	25	0.5	0.399	0.223	0.000	0.000	0.043	0.000	0.000	0.914	0.416	0.181
100	50		0.570	0.274	0.000	0.000	0.033	0.000	0.000	0.988	0.503	0.165
100	100		0.705	0.434	0.000	0.000	0.059	0.000	0.000	0.998	0.596	0.164
300	10		0.478	0.996	0.709	0.350	0.979	0.627	0.263	0.999	0.953	0.794
300	25	0.5	0.740	1.000	0.878	0.451	1.000	0.817	0.327	1.000	1.000	0.957
300	50		0.925	1.000	0.985	0.601	1.000	0.955	0.398	1.000	1.000	0.998
300	100		0.990	1.000	0.999	0.814	1.000	0.997	0.557	1.000	1.000	1.000
500	10		0.547	1.000	0.976	0.840	1.000	0.971	0.802	1.000	0.996	0.964
500	25	0.5	0.819	1.000	1.000	0.977	1.000	1.000	0.962	1.000	1.000	1.000
500	50		0.963	1.000	1.000	1.000	1.000	1.000	0.998	1.000	1.000	1.000
500	100		0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 4: Empirical size and power of panel tests
(cross correlation, strong serial correlation)

			constant model									
T	N	N_1/N	LM	$\hat{S}_K^{ols}(1)$	$\hat{S}_K^{ols}(2)$	$\hat{S}_K^{ols}(3)$	$\hat{S}_K(1)$	$\hat{S}_K(2)$	$\hat{S}_K(3)$	$\tilde{S}_K(1)$	$\tilde{S}_K(2)$	$\tilde{S}_K(3)$
100	10		0.052	0.025	0.003	0.001	0.005	0.001	0.001	0.233	0.092	0.076
100	25	0	0.055	0.043	0.002	0.000	0.002	0.000	0.000	0.402	0.078	0.055
100	50		0.054	0.064	0.000	0.000	0.002	0.000	0.000	0.582	0.067	0.036
100	100		0.024	0.019	0.000	0.000	0.000	0.000	0.000	0.457	0.030	0.018
300	10		0.055	0.020	0.007	0.003	0.009	0.005	0.003	0.123	0.070	0.072
300	25	0	0.040	0.019	0.001	0.001	0.010	0.001	0.000	0.241	0.072	0.049
300	50		0.013	0.018	0.000	0.000	0.008	0.000	0.000	0.419	0.060	0.042
300	100		0.004	0.002	0.000	0.000	0.000	0.000	0.000	0.290	0.044	0.036
500	10		0.032	0.020	0.014	0.009	0.012	0.006	0.007	0.090	0.058	0.058
500	25	0	0.019	0.008	0.002	0.001	0.010	0.001	0.001	0.152	0.059	0.051
500	50		0.007	0.005	0.000	0.000	0.006	0.000	0.000	0.249	0.053	0.044
500	100		0.001	0.002	0.000	0.000	0.000	0.000	0.000	0.161	0.035	0.038
100	10		0.142	0.250	0.024	0.011	0.070	0.005	0.002	0.671	0.296	0.184
100	25	0.2	0.182	0.478	0.015	0.002	0.106	0.001	0.000	0.911	0.392	0.201
100	50		0.262	0.711	0.008	0.000	0.144	0.000	0.000	0.984	0.451	0.179
100	100		0.280	0.739	0.001	0.000	0.088	0.000	0.000	0.993	0.437	0.163
300	10		0.216	0.705	0.366	0.223	0.633	0.289	0.176	0.937	0.739	0.590
300	25	0.2	0.323	0.960	0.568	0.295	0.930	0.467	0.239	0.999	0.948	0.838
300	50		0.432	0.999	0.753	0.353	0.994	0.603	0.251	1.000	0.995	0.957
300	100		0.571	1.000	0.936	0.618	1.000	0.835	0.430	1.000	1.000	0.996
500	10		0.248	0.935	0.721	0.564	0.921	0.694	0.539	0.989	0.913	0.815
500	25	0.2	0.400	0.999	0.947	0.826	0.999	0.925	0.774	1.000	0.998	0.979
500	50		0.552	1.000	1.000	0.963	1.000	0.993	0.915	1.000	1.000	0.999
500	100		0.715	1.000	1.000	0.998	1.000	1.000	0.996	1.000	1.000	1.000
100	10		0.318	0.721	0.163	0.042	0.375	0.045	0.009	0.944	0.636	0.406
100	25	0.5	0.476	0.947	0.218	0.022	0.652	0.029	0.001	0.999	0.822	0.500
100	50		0.731	0.999	0.322	0.016	0.909	0.020	0.001	1.000	0.957	0.629
100	100		0.874	1.000	0.416	0.008	0.969	0.015	0.000	1.000	0.990	0.721
300	10		0.449	0.999	0.944	0.819	0.999	0.926	0.753	1.000	0.995	0.968
300	25	0.5	0.720	1.000	1.000	0.957	1.000	0.995	0.928	1.000	1.000	0.999
300	50		0.913	1.000	1.000	1.000	1.000	1.000	0.995	1.000	1.000	1.000
300	100		0.989	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
500	10		0.515	1.000	0.998	0.986	1.000	0.998	0.984	1.000	1.000	0.998
500	25	0.5	0.778	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
500	50		0.949	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
500	100		0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 4: (continued)
(cross correlation, strong serial correlation)

			trend model									
T	N	N_1/N	LM	$\hat{S}_K^{ols}(1)$	$\hat{S}_K^{ols}(2)$	$\hat{S}_K^{ols}(3)$	$\hat{S}_K(1)$	$\hat{S}_K(2)$	$\hat{S}_K(3)$	$\tilde{S}_K(1)$	$\tilde{S}_K(2)$	$\tilde{S}_K(3)$
100	10		0.054	0.000	0.000	0.000	0.000	0.000	0.000	0.092	0.045	0.067
100	25	0	0.059	0.000	0.000	0.000	0.000	0.000	0.000	0.093	0.020	0.031
100	50		0.046	0.000	0.000	0.000	0.000	0.000	0.000	0.105	0.008	0.015
100	100		0.013	0.000	0.000	0.000	0.000	0.000	0.000	0.036	0.002	0.010
300	10		0.051	0.003	0.000	0.000	0.001	0.001	0.001	0.093	0.059	0.071
300	25	0	0.031	0.000	0.000	0.001	0.000	0.000	0.000	0.134	0.045	0.039
300	50		0.012	0.000	0.000	0.000	0.000	0.000	0.000	0.211	0.025	0.020
300	100		0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.107	0.018	0.021
500	10		0.043	0.005	0.004	0.003	0.003	0.001	0.001	0.078	0.057	0.060
500	25	0	0.008	0.000	0.001	0.000	0.000	0.000	0.000	0.105	0.047	0.045
500	50		0.006	0.000	0.000	0.000	0.000	0.000	0.000	0.153	0.034	0.029
500	100		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.077	0.023	0.026
100	10		0.106	0.010	0.000	0.000	0.001	0.000	0.000	0.295	0.076	0.070
100	25	0.2	0.144	0.001	0.000	0.000	0.000	0.000	0.000	0.384	0.049	0.031
100	50		0.144	0.001	0.000	0.000	0.000	0.000	0.000	0.450	0.023	0.012
100	100		0.116	0.000	0.000	0.000	0.000	0.000	0.000	0.346	0.009	0.010
300	10		0.174	0.269	0.048	0.023	0.199	0.035	0.016	0.796	0.455	0.312
300	25	0.2	0.256	0.473	0.031	0.005	0.314	0.018	0.004	0.978	0.675	0.434
300	50		0.314	0.657	0.007	0.001	0.433	0.006	0.001	1.000	0.822	0.504
300	100		0.341	0.760	0.007	0.000	0.506	0.003	0.000	1.000	0.946	0.691
500	10		0.207	0.682	0.312	0.155	0.626	0.260	0.135	0.943	0.730	0.565
500	25	0.2	0.310	0.940	0.475	0.199	0.898	0.375	0.140	1.000	0.951	0.812
500	50		0.413	0.996	0.659	0.232	0.988	0.485	0.126	1.000	0.996	0.934
500	100		0.516	1.000	0.814	0.230	1.000	0.755	0.229	1.000	1.000	0.995
100	10		0.207	0.054	0.000	0.000	0.008	0.000	0.000	0.597	0.155	0.081
100	25	0.5	0.328	0.050	0.000	0.000	0.002	0.000	0.000	0.794	0.113	0.032
100	50		0.459	0.062	0.000	0.000	0.000	0.000	0.000	0.930	0.091	0.012
100	100		0.564	0.022	0.000	0.000	0.000	0.000	0.000	0.962	0.048	0.006
300	10		0.457	0.928	0.424	0.143	0.889	0.346	0.108	0.998	0.903	0.702
300	25	0.5	0.704	0.999	0.551	0.114	0.993	0.413	0.056	1.000	0.993	0.881
300	50		0.898	1.000	0.794	0.097	1.000	0.591	0.042	1.000	1.000	0.981
300	100		0.976	1.000	0.944	0.140	1.000	0.794	0.037	1.000	1.000	1.000
500	10		0.515	1.000	0.936	0.732	0.998	0.910	0.655	1.000	0.994	0.947
500	25	0.5	0.786	1.000	0.996	0.882	1.000	0.993	0.830	1.000	1.000	0.998
500	50		0.946	1.000	1.000	0.993	1.000	1.000	0.968	1.000	1.000	1.000
500	100		0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 5: Empirical size and power of panel tests
(cross cointegration, cross correlation, diversified serial correlation)

			constant model									
T	N	N_1/N	LM	$\hat{S}_K^{ols}(1)$	$\hat{S}_K^{ols}(2)$	$\hat{S}_K^{ols}(3)$	$\hat{S}_K(1)$	$\hat{S}_K(2)$	$\hat{S}_K(3)$	$\tilde{S}_K(1)$	$\tilde{S}_K(2)$	$\tilde{S}_K(3)$
100	10		0.058	0.010	0.018	0.019	0.010	0.009	0.013	0.065	0.068	0.084
100	25	0	0.078	0.010	0.009	0.014	0.005	0.007	0.011	0.070	0.082	0.103
100	50		0.068	0.014	0.009	0.010	0.003	0.005	0.005	0.074	0.090	0.117
100	100		0.072	0.003	0.005	0.008	0.001	0.003	0.004	0.075	0.101	0.130
300	10		0.052	0.020	0.026	0.021	0.029	0.018	0.022	0.059	0.047	0.066
300	25	0	0.046	0.031	0.031	0.020	0.015	0.017	0.017	0.061	0.067	0.060
300	50		0.048	0.015	0.012	0.017	0.011	0.011	0.015	0.060	0.064	0.071
300	100		0.050	0.016	0.010	0.015	0.008	0.007	0.009	0.064	0.073	0.075
500	10		0.050	0.046	0.031	0.034	0.026	0.027	0.033	0.055	0.053	0.065
500	25	0	0.056	0.020	0.028	0.036	0.018	0.020	0.020	0.051	0.054	0.059
500	50		0.052	0.028	0.018	0.026	0.016	0.016	0.016	0.062	0.057	0.066
500	100		0.057	0.023	0.021	0.023	0.009	0.012	0.011	0.057	0.063	0.072
100	10		0.159	0.179	0.075	0.038	0.123	0.052	0.031	0.529	0.372	0.295
100	25	0.2	0.225	0.293	0.102	0.072	0.174	0.069	0.039	0.732	0.519	0.391
100	50		0.252	0.359	0.137	0.066	0.197	0.072	0.036	0.838	0.589	0.455
100	100		0.259	0.341	0.127	0.051	0.209	0.072	0.034	0.902	0.633	0.495
300	10		0.258	0.858	0.659	0.501	0.833	0.603	0.453	0.951	0.824	0.711
300	25	0.2	0.313	0.972	0.814	0.632	0.979	0.825	0.634	0.999	0.974	0.902
300	50		0.362	0.997	0.921	0.746	0.999	0.915	0.735	1.000	0.996	0.964
300	100		0.398	0.998	0.954	0.789	1.000	0.970	0.807	1.000	1.000	0.992
500	10		0.272	0.981	0.906	0.817	0.985	0.903	0.780	0.997	0.962	0.899
500	25	0.2	0.374	1.000	0.994	0.948	1.000	0.994	0.957	1.000	0.999	0.994
500	50		0.437	1.000	1.000	0.984	1.000	1.000	0.993	1.000	1.000	1.000
500	100		0.472	1.000	1.000	0.997	1.000	1.000	0.999	1.000	1.000	1.000
100	10		0.315	0.614	0.237	0.118	0.403	0.132	0.061	0.920	0.673	0.495
100	25	0.5	0.400	0.798	0.331	0.160	0.547	0.179	0.087	0.987	0.797	0.575
100	50		0.400	0.866	0.375	0.193	0.622	0.196	0.087	0.997	0.840	0.611
100	100		0.389	0.887	0.345	0.195	0.651	0.197	0.096	1.000	0.885	0.639
300	10		0.387	0.999	0.951	0.855	0.998	0.932	0.781	1.000	0.993	0.957
300	25	0.5	0.445	1.000	0.994	0.938	1.000	0.994	0.920	1.000	1.000	0.997
300	50		0.494	1.000	1.000	0.975	1.000	0.999	0.966	1.000	1.000	0.999
300	100		0.499	1.000	1.000	0.993	1.000	1.000	0.990	1.000	1.000	1.000
500	10		0.427	1.000	0.998	0.988	1.000	0.999	0.984	1.000	1.000	0.997
500	25	0.5	0.486	1.000	1.000	0.999	1.000	1.000	0.999	1.000	1.000	1.000
500	50		0.547	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
500	100		0.596	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 5: (continued)
(cross cointegration, cross correlation, diversified serial correlation)

			trend model									
T	N	N_1/N	LM	$\hat{S}_K^{ols}(1)$	$\hat{S}_K^{ols}(2)$	$\hat{S}_K^{ols}(3)$	$\hat{S}_K(1)$	$\hat{S}_K(2)$	$\hat{S}_K(3)$	$\tilde{S}_K(1)$	$\tilde{S}_K(2)$	$\tilde{S}_K(3)$
100	10		0.053	0.009	0.008	0.015	0.005	0.006	0.009	0.066	0.078	0.099
100	25	0	0.049	0.008	0.005	0.017	0.003	0.003	0.007	0.070	0.094	0.132
100	50		0.063	0.001	0.004	0.005	0.001	0.002	0.003	0.067	0.107	0.139
100	100		0.060	0.001	0.003	0.004	0.001	0.001	0.002	0.064	0.103	0.164
300	10		0.053	0.010	0.014	0.022	0.021	0.015	0.014	0.061	0.055	0.073
300	25	0	0.047	0.019	0.016	0.020	0.009	0.011	0.012	0.066	0.076	0.077
300	50		0.065	0.013	0.012	0.005	0.004	0.006	0.009	0.067	0.072	0.089
300	100		0.060	0.007	0.007	0.013	0.002	0.003	0.004	0.069	0.087	0.098
500	10		0.057	0.040	0.027	0.030	0.021	0.020	0.028	0.057	0.057	0.070
500	25	0	0.048	0.016	0.023	0.028	0.012	0.013	0.015	0.054	0.060	0.065
500	50		0.049	0.024	0.013	0.020	0.009	0.010	0.011	0.065	0.064	0.077
500	100		0.064	0.015	0.015	0.015	0.004	0.007	0.006	0.063	0.073	0.084
100	10		0.132	0.035	0.007	0.006	0.021	0.004	0.003	0.324	0.191	0.152
100	25	0.2	0.162	0.039	0.002	0.001	0.017	0.001	0.001	0.438	0.251	0.188
100	50		0.188	0.033	0.001	0.000	0.014	0.002	0.001	0.511	0.272	0.195
100	100		0.224	0.033	0.002	0.000	0.008	0.000	0.000	0.559	0.291	0.208
300	10		0.228	0.556	0.263	0.152	0.512	0.223	0.118	0.856	0.601	0.432
300	25	0.2	0.307	0.796	0.368	0.158	0.782	0.329	0.137	0.985	0.835	0.630
300	50		0.374	0.904	0.431	0.188	0.907	0.410	0.165	1.000	0.935	0.753
300	100		0.455	0.947	0.469	0.170	0.962	0.475	0.169	1.000	0.978	0.850
500	10		0.267	0.914	0.675	0.466	0.893	0.621	0.414	0.980	0.854	0.705
500	25	0.2	0.395	0.994	0.864	0.617	0.995	0.866	0.599	1.000	0.989	0.914
500	50		0.460	1.000	0.949	0.737	1.000	0.955	0.724	1.000	0.999	0.977
500	100		0.533	1.000	0.980	0.813	1.000	0.990	0.811	1.000	1.000	0.998
100	10		0.259	0.136	0.004	0.001	0.055	0.002	0.000	0.666	0.321	0.182
100	25	0.5	0.367	0.198	0.007	0.000	0.071	0.000	0.000	0.805	0.394	0.213
100	50		0.408	0.197	0.004	0.000	0.071	0.001	0.000	0.874	0.425	0.226
100	100		0.442	0.216	0.000	0.000	0.073	0.000	0.000	0.898	0.448	0.245
300	10		0.461	0.967	0.589	0.302	0.950	0.510	0.232	0.999	0.919	0.724
300	25	0.5	0.582	0.999	0.788	0.405	0.997	0.702	0.316	1.000	0.991	0.880
300	50		0.650	1.000	0.889	0.455	1.000	0.789	0.379	1.000	1.000	0.951
300	100		0.680	1.000	0.906	0.501	1.000	0.845	0.417	1.000	1.000	0.976
500	10		0.513	1.000	0.961	0.795	0.999	0.950	0.749	1.000	0.995	0.951
500	25	0.5	0.656	1.000	0.997	0.927	1.000	0.999	0.908	1.000	1.000	0.998
500	50		0.702	1.000	1.000	0.982	1.000	1.000	0.967	1.000	1.000	1.000
500	100		0.739	1.000	1.000	0.995	1.000	1.000	0.990	1.000	1.000	1.000

6. Conclusion

In this paper we have proposed tests assuming a null hypothesis of cointegration. Contrary to the single equation cointegration tests in the literature where the limiting distributions are non-standard, we show that our tests have a standard normal asymptotic distribution. Our tests are transposed to panel data cointegration tests allowing for cross-section dependence and serial correlation. We have proved for fixed N and $T \to \infty$ that the limiting distributions of our statistics are standard normals and proposed a bias correction which is shown to work well in finite samples via Monte Carlo simulations, particularly when T is larger than N. Finally, our tests are robust to the likely presence of cointegration across units which is often the case in macroeconomic data.

Appendix

In this appendix, $\bar{c}$ signifies a generic positive constant that may differ from place to place.

Proof of Lemma 1

Using expression (5), we decompose $\eta_t \eta_{t-K}$ into 5 parts as follows:

$$\begin{aligned}
\eta_t \eta_{t-K} &= \sum_{j=-\infty}^{\infty} \phi_j \xi_{t-j} \sum_{\ell=-\infty}^{\infty} \phi_\ell \xi_{t-K-\ell} \\
&= \sum_{j=1}^{K-1} \sum_{\ell=0}^{\infty} g_t(j,\ell) + \sum_{j=1}^{K-1} \sum_{\ell=1-K}^{-1} g_t(j,\ell) + \sum_{j=1-K}^{0} \sum_{\ell=1-K}^{\infty} g_t(j,\ell) \\
&\qquad + \sum_{|j|\geq K}^{\infty} \sum_{\ell=-\infty}^{\infty} g_t(j,\ell) + \sum_{j=1-K}^{K-1} \sum_{\ell=-\infty}^{-K} g_t(j,\ell) \\
&\equiv C_{1,t} + C_{2,t} + C_{3,t} + r_{2,t} + r_{3,t}, \quad \text{say,} \qquad (21)
\end{aligned}$$

where $g_t(j,\ell) = \phi_j \phi_\ell \xi_{t-j} \xi_{t-K-\ell}$.

For $C_{1,t}$, we can see that

$$C_{1,t} = \sum_{j=1}^{K-1} \sum_{\ell=0}^{j-1} g_t(j,\ell) + \sum_{j=1}^{K-1} \sum_{\ell=j}^{\infty} g_t(j,\ell).$$

The first term becomes

$$
\begin{aligned}
\sum_{j=1}^{K-1}\sum_{\ell=0}^{j-1} g_t(j,\ell) &= [g_t(K-1,0)] + [g_t(K-2,0) + g_t(K-1,1)] \\
&\quad + \cdots + [g_t(1,0) + g_t(2,1) + \cdots + g_t(K-1,K-2)] \\
&= \sum_{j=1}^{K-1}\sum_{\ell=0}^{j-1} g_t(K-j+\ell,\ell) \\
&= \sum_{j=1}^{K-1}\sum_{\ell=K-j}^{K-1} g_t(\ell, j+\ell-K) \\
&= \sum_{j=1}^{K-1}\sum_{\ell=K-j}^{K-1} \phi_\ell\phi_{j+\ell-K}\xi_{t-\ell}\xi_{t-\ell-j}, \qquad (22)
\end{aligned}
$$

where the third equality holds by re-defining ℓ as $K-j+\ell$. Similarly, we have

$$
\begin{aligned}
\sum_{j=1}^{K-1}\sum_{\ell=j}^{\infty} g_t(j,\ell) &= \sum_{j=0}^{\infty}\sum_{\ell=1}^{K-1} g_t(\ell, j+\ell) \\
&= \sum_{j=K}^{\infty}\sum_{\ell=1}^{K-1} g_t(\ell, j+\ell-K) \\
&= \sum_{j=K}^{\infty}\sum_{\ell=1}^{K-1} \phi_\ell\phi_{j+\ell-K}\xi_{t-\ell}\xi_{t-\ell-j}, \qquad (23)
\end{aligned}
$$

where the second equality is obtained by re-defining j as $j+K$.

Similarly, we have

$$
\begin{aligned}
C_{2,t} &= \sum_{j=1}^{K-1}\sum_{\ell=1-K}^{-1} g_t(j,\ell) \\
&= \sum_{j=1}^{K-1}\sum_{\ell=1}^{K-1} \phi_j\phi_{\ell-K}\xi_{t-j}\xi_{t-\ell}
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{j=1}^{K-1} \phi_j\phi_{j-K}\xi_{t-j}^2 + \sum_{j=1}^{K-2}\sum_{\ell=j+1}^{K-1} (\phi_j\phi_{\ell-K} + \phi_\ell\phi_{j-K})\xi_{t-j}\xi_{t-\ell} \\
&= r_{1,t} + \sum_{j=1}^{K-2}\sum_{\ell=1}^{K-j-1} (\phi_\ell\phi_{j+\ell-K} + \phi_{j+\ell}\phi_{\ell-K})\xi_{t-\ell}\xi_{t-\ell-j}. \qquad (24)
\end{aligned}
$$

Then, we have, from (22)–(24),

$$
C_{1,t} + C_{2,t} = r_{1,t} + \sum_{j=1}^{\infty}\sum_{\ell=1}^{K-1} \phi_\ell\phi_{j+\ell-K}\xi_{t-\ell}\xi_{t-\ell-j} + \sum_{j=1}^{K-2}\sum_{\ell=1}^{K-j-1} \phi_{j+\ell}\phi_{\ell-K}\xi_{t-\ell}\xi_{t-\ell-j}. \qquad (25)
$$

For $C_{3,t}$,

$$\begin{aligned}
C_{3,t} &= \sum_{j=1-K}^{0}\sum_{\ell=1-K}^{\infty} g_t(j,\ell) \\
&= \sum_{j=1-K}^{0}\sum_{\ell=1-K}^{j} g_t(j,\ell) + \sum_{j=1-K}^{0}\sum_{\ell=j+1}^{\infty} g_t(j,\ell) \\
&= \sum_{j=1}^{K}\sum_{\ell=1}^{j} g_t(-j+\ell,\ell-K) + \sum_{j=1}^{\infty}\sum_{\ell=1}^{K} g_t(\ell-K,j+\ell-K) \\
&= \sum_{j=1}^{K}\sum_{\ell=1-j}^{0} g_t(\ell,j+\ell-K) + \sum_{j=1}^{\infty}\sum_{\ell=1-K}^{0} g_t(\ell,j+\ell) \\
&= \sum_{j=1}^{K}\sum_{\ell=1-j}^{0} g_t(\ell,j+\ell-K) + \sum_{j=K+1}^{\infty}\sum_{\ell=1-K}^{0} g_t(\ell,j+\ell-K) \\
&= \sum_{j=1}^{\infty}\sum_{\ell=(1-j)\vee(1-K)}^{0} g_t(\ell,j+\ell-K) \\
&= \sum_{j=1}^{\infty}\sum_{\ell=(1-j)\vee(1-K)}^{0} \phi_\ell\phi_{j+\ell-K}\xi_{t-\ell}\xi_{t-\ell-j}, \qquad (26)
\end{aligned}$$

and then from (25) and (26), we have

$$\begin{aligned}
&C_{1,t}+C_{2,t}+C_{3,t} \\
&= r_{1,t} + \sum_{j=1}^{\infty}\sum_{\ell=(1-j)\vee(1-K)}^{K-1} \phi_\ell\phi_{j+\ell-K}\xi_{t-\ell}\xi_{t-\ell-j} + \sum_{j=1}^{K-2}\sum_{\ell=1}^{K-j-1} \phi_{j+\ell}\phi_{\ell-K}\xi_{t-\ell}\xi_{t-\ell-j} \\
&= r_{1,t} + C_{a,t} + C_{b,t}, \quad \text{say.} \qquad (27)
\end{aligned}$$

We next apply the B–N decomposition to $C_{a,t}$ and $C_{b,t}$. For $C_{a,t}$, we consider three cases where $\ell = 0$, $\ell \geq 1$ and $\ell \leq -1$. For $\ell = 0$, we have

$$C_{a,t} = \sum_{j=1}^{\infty} \phi_0\phi_{j-K}\xi_t\xi_{t-j}, \qquad (28)$$

while for $\ell \geq 1$,

$$\begin{aligned}
C_{a,t} &= \sum_{j=1}^{\infty}\sum_{\ell=1}^{K-1}\phi_\ell\phi_{j+\ell-K}L^\ell\xi_t\xi_{t-j} \\
&= \sum_{j=1}^{\infty}\sum_{\ell=1}^{K-1}\phi_\ell\phi_{j+\ell-K}[1-(1-L^\ell)]\xi_t\xi_{t-j} \\
&= \sum_{j=1}^{\infty}\sum_{\ell=1}^{K-1}\phi_\ell\phi_{j+\ell-K}\xi_t\xi_{t-j} - \Delta\sum_{j=1}^{\infty}\sum_{\ell=1}^{K-1}\phi_\ell\phi_{j+\ell-K}\sum_{i=0}^{\ell-1}L^i\xi_t\xi_{t-j} \\
&= \sum_{j=1}^{\infty}\sum_{\ell=1}^{K-1}\phi_\ell\phi_{j+\ell-K}\xi_t\xi_{t-j} - \Delta\sum_{j=1}^{\infty}\sum_{i=0}^{K-2}\left(\sum_{\ell=i+1}^{K-1}\phi_\ell\phi_{j+\ell-K}\right)L^i\xi_t\xi_{t-j} \\
&= \sum_{j=1}^{\infty}\sum_{\ell=1}^{K-1}\phi_\ell\phi_{j+\ell-K}\xi_t\xi_{t-j} - \Delta\tilde{r}_{1,t}. \qquad (29)
\end{aligned}$$

For $\ell \leq -1$, it is sufficient to consider the case where $j \geq 2$. For $j = 2, \cdots, K-1$, the summand of $C_{a,t}$ becomes

$$\begin{aligned}
&\sum_{\ell=1-j}^{-1}\phi_\ell\phi_{j+\ell-K}L^\ell\xi_t\xi_{t-j} \\
&\quad= \sum_{\ell=1-j}^{-1}\phi_\ell\phi_{j+\ell-K}[1-(1-L^\ell)]\xi_t\xi_{t-j} \\
&\quad= \sum_{\ell=1-j}^{-1}\phi_\ell\phi_{j+\ell-K}\xi_t\xi_{t-j} - \Delta^+\sum_{\ell=1-j}^{-1}\phi_\ell\phi_{j+\ell-K}\sum_{i=\ell+1}^{0}L^i\xi_t\xi_{t-j} \\
&\quad= \sum_{\ell=1-j}^{-1}\phi_\ell\phi_{j+\ell-K}\xi_t\xi_{t-j} - \Delta^+\sum_{i=2-j}^{0}\left(\sum_{\ell=1-j}^{i-1}\phi_\ell\phi_{j+\ell-K}\right)L^i\xi_t\xi_{t-j}, \qquad (30)
\end{aligned}$$

where $\Delta^+ = (1-L^{-1})$ and we used the relation $(1-L^\ell) = (1-L^{-1})(1+L^{-1}+\cdots+L^{\ell+1})$ for $\ell < 0$, while for $j \geq K$, it can be expressed as

$$\begin{aligned}
\sum_{\ell=1-K}^{-1}\phi_\ell\phi_{j+\ell-K}L^\ell\xi_t\xi_{t-j} &= \sum_{\ell=1-K}^{-1}\phi_\ell\phi_{j+\ell-K}\xi_t\xi_{t-j} - \Delta^+\sum_{\ell=1-K}^{-1}\phi_\ell\phi_{j+\ell-K}\sum_{i=\ell+1}^{0}L^i\xi_t\xi_{t-j} \\
&= \sum_{\ell=1-K}^{-1}\phi_\ell\phi_{j+\ell-K}\xi_t\xi_{t-j} - \Delta^+\sum_{i=2-K}^{0}\left(\sum_{\ell=1-K}^{i-1}\phi_\ell\phi_{j+\ell-K}\right)L^i\xi_t\xi_{t-j}. \qquad (31)
\end{aligned}$$

From (30) and (31), $C_{a,t}$ for $\ell \leq -1$ becomes

$$\begin{aligned}
C_{a,t} &= \sum_{j=2}^{\infty}\sum_{\ell=1-(j\wedge K)}^{-1}\phi_\ell\phi_{j+\ell-K}\xi_t\xi_{t-j} - \Delta^+\sum_{j=2}^{\infty}\sum_{i=2-(j\wedge K)}^{0}\left(\sum_{\ell=1-(j\wedge K)}^{i-1}\phi_\ell\phi_{j+\ell-K}\right)L^i\xi_t\xi_{t-j} \\
&= \sum_{j=2}^{\infty}\sum_{\ell=1-(j\wedge K)}^{-1}\phi_\ell\phi_{j+\ell-K}\xi_t\xi_{t-j} - \Delta^+\tilde{r}_t^+. \qquad (32)
\end{aligned}$$

Combining (28), (29) and (32), we can see that

$$
\begin{aligned}
C_{a,t} &= \sum_{j=1}^{\infty} \sum_{\ell=1-(j\wedge K)}^{K-1} \phi_\ell \phi_{j+\ell-K} \xi_t \xi_{t-j} - \Delta \tilde{r}_{1,t} - \Delta^+ \tilde{r}_t^+ \\
&= \sum_{j=1}^{\infty} G_{1,j} \xi_t \xi_{t-j} - \Delta \tilde{r}_{1,t} - \Delta^+ \tilde{r}_t^+. \qquad (33)
\end{aligned}
$$

In exactly the same way, we have

$$
\begin{aligned}
C_{b,t} &= \sum_{j=1}^{K-2} \sum_{\ell=1}^{K-j-1} \phi_{j+\ell} \phi_{\ell-K} \xi_t \xi_{t-j} - \Delta \sum_{j=1}^{K-2} \sum_{\ell=1}^{K-j-1} \phi_{j+\ell} \phi_{\ell-K} \sum_{i=0}^{\ell-1} L^i \xi_t \xi_{t-j} \\
&= \sum_{j=1}^{K-2} G_{2,j} \xi_t \xi_{t-j} - \Delta \sum_{j=1}^{K-2} \sum_{i=0}^{K-j-2} \left(\sum_{\ell=i+1}^{K-j-1} \phi_{j+\ell} \phi_{\ell-K} \right) L^i \xi_t \xi_{t-j} \\
&= \sum_{j=1}^{K-2} G_{2,j} \xi_t \xi_{t-j} - \Delta \tilde{r}_{2,t}. \qquad (34)
\end{aligned}
$$

Combining (21), (27), (33) and (34), we obtain (10).■

Proof of Lemma 2

From (10) in Lemma 1, we can see that

$$
\begin{aligned}
\frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} \eta_t \eta_{t-K} &= \frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} \sum_{j=1}^{\infty} G_j \xi_t \xi_{t-j} + \frac{1}{\sqrt{T}} \left(\tilde{r}_0 - \tilde{r}_{[Tr]} - r_1^+ + r_{[Tr]}^+ \right) \\
&\quad + \frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} (r_{1,t} + r_{2,t} + r_{3,t}). \qquad (35)
\end{aligned}
$$

We will show that the FCLT holds for the first term on the right-hand side while the other terms are negligible, using the following lemma:

Lemma A.1 *For $\{\phi_j\}_{j=-\infty}^{\infty}$ satisfying the condition given by* (5), *(i)* $\sum_{|j|\geq K}^{\infty} |\phi_j| = o\left(\frac{1}{K^2}\right)$ *and* $\sum_{|j|\geq K}^{\infty} |\phi_j|^2 = o\left(\frac{1}{K^4}\right)$, *(ii)* $\sum_{j=1}^{\infty} |G_j| < \infty$, *(iii)* $\sum_{j=1}^{\infty} \sum_{\ell=0}^{K-2} |\tilde{G}_{1,\ell}| < \infty$, *(iv)* $\sum_{j=1}^{K-2} \sum_{\ell=0}^{K-j-2} |\tilde{G}_{2,\ell}| < \infty$, *(v)* $\sum_{j=2}^{\infty} \sum_{\ell=2-(j\wedge K)}^{0} |\tilde{G}_\ell^+| < \infty$. *The relations (ii)–(v) hold uniformly over K.*

Proof of Lemma A.1: (i) is shown by

$$
\begin{aligned}
\sum_{|j|\geq K}^{\infty} |\phi_j| &\leq \frac{1}{K^2} \sum_{|j|\geq K}^{\infty} |j|^2 |\phi_j| = o\left(\frac{1}{K^2}\right), \\
\sum_{|j|\geq K}^{\infty} |\phi_j|^2 &\leq \frac{1}{K^4} \sum_{|j|\geq K}^{\infty} |j|^4 |\phi_j|^2 = o\left(\frac{1}{K^4}\right).
\end{aligned}
$$

For (ii)-(v), we have

$$
\begin{aligned}
\sum_{j=1}^{\infty}|G_j| &\le \sum_{j=1}^{\infty}|G_{1,j}| + \sum_{j=1}^{K-2}|G_{2j}| \\
&\le \sum_{j=1}^{\infty}\sum_{\ell=1-(j\wedge K)}^{K-1}|\phi_\ell||\phi_{j+\ell-K}| + \sum_{j=1}^{K-2}\sum_{\ell=1}^{K-j-1}|\phi_{j+\ell}||\phi_{\ell-K}| \\
&\le \sum_{\ell=1-K}^{K-1}|\phi_\ell|\sum_{j=1}^{\infty}|\phi_{j+\ell-K}| + \sum_{\ell=1}^{K-2}|\phi_{\ell-K}|\sum_{j=1}^{K-2}|\phi_{j+\ell}| \\
&\le \left(\sum_{\ell=-\infty}^{\infty}|\phi_\ell|\right)^2 + \sum_{\ell=-\infty}^{\infty}|\phi_\ell|\sum_{j=1}^{\infty}|\phi_j| < \infty.
\end{aligned}
$$

$$
\begin{aligned}
\sum_{j=1}^{\infty}\sum_{\ell=0}^{K-2}|\tilde{G}_{1,\ell}| &\le \sum_{j=1}^{\infty}\sum_{\ell=0}^{K-2}\sum_{i=\ell+1}^{K-1}|\phi_i||\phi_{i+j-K}| \\
&\le \sum_{\ell=0}^{K-2}\sum_{i=\ell+1}^{K-1}|\phi_i|\sum_{j=-\infty}^{\infty}|\phi_j| \\
&= \sum_{i=1}^{K-1} i|\phi_i|\sum_{j=-\infty}^{\infty}|\phi_j| < \infty.
\end{aligned}
$$

$$
\begin{aligned}
\sum_{j=1}^{K-2}\sum_{\ell=0}^{K-j-2}|\tilde{G}_{2,\ell}| &\le \sum_{j=1}^{K-2}\sum_{\ell=0}^{K-j-2}\sum_{i=\ell+1}^{K-j-1}|\phi_{i+j}||\phi_{i-K}| \\
&= \sum_{j=1}^{K-2}\sum_{i=1}^{K-j-1} i|\phi_{i+j}||\phi_{i-K}| \\
&\le \sum_{i=1}^{K-2}|\phi_{i-K}|\sum_{j=1}^{K-2}(i+j)|\phi_{i+j}| \\
&\le \sum_{i=-\infty}^{\infty}|\phi_i|\sum_{j=1}^{\infty} j|\phi_j| < \infty.
\end{aligned}
$$

$$
\begin{aligned}
\sum_{j=2}^{\infty}\sum_{\ell=2-(j\wedge K)}^{0}|\tilde{G}_\ell^+| &\le \sum_{j=2}^{\infty}\sum_{\ell=2-(j\wedge K)}^{0}\sum_{i=1-(j\wedge K)}^{\ell-1}|\phi_i||\phi_{i+j-K}| \\
&\le \sum_{\ell=2-K}^{0}\sum_{i=1-K}^{\ell-1}|\phi_i|\sum_{j=2}^{\infty}|\phi_{i+j-K}| \\
&\le \sum_{i=1-K}^{-1}|i||\phi_i|\sum_{j=-\infty}^{\infty}|\phi_j| < \infty. \blacksquare
\end{aligned}
$$

Note that the absolute summability in Lemma A.1(ii)–(v) implies the square summability of the corresponding terms. Using Lemma A.1, we show that all the term on the right hand side of (35), except for the first term, are negligible.

Lemma A.2 *For $\tilde{r}_t$, $\tilde{r}_t^+$, $r_{1,t}$, $r_{2,t}$ and $r_{3,t}$ in* (35)*, (i)* $\sup_{0\le r\le 1}\left|\frac{1}{\sqrt{T}}\tilde{r}_{[Tr]}\right| = o_p(1)$ *and* $\sup_{0\le r\le 1}\left|\frac{1}{\sqrt{T}}\tilde{r}^+_{[Tr]}\right| = o_p(1)$*. (ii)* $\sup_{0\le r\le 1}\left|\frac{1}{\sqrt{T}}\sum_{t=1}^{[Tr]} r_{i,t}\right| = o_p(1)$ *for $i = 1, 2$ and* 3.

Proof of Lemma A.2: (i) We first note that $\tilde{r}_t = \tilde{r}_{1,t} + \tilde{r}_{2,t}$ as defined in Lemma 1. Since

$$P\left(\sup_{0\le r\le 1}\left|\frac{1}{\sqrt{T}}\tilde{r}_{i,t}\right| \ge \varepsilon\right) \le TP\left(\frac{1}{\sqrt{T}}\left|\tilde{r}_{i,t}\right| \ge \varepsilon\right) \le \frac{1}{\varepsilon^4 T}E[\tilde{r}^4_{i,t}]$$

for $i = 1$ and 2, it is sufficient to prove that $E[\tilde{r}^4_{i,t}] < \infty$ for $i = 1$ and 2. Noting that non-zero terms of $E[\tilde{r}^4_{i,t}]$ are related to the products among $E[\xi_t^2]$, $E[\xi_t^3]$ and $E[\xi_t^4]$, all of which are bounded by assumption, we can see that

$$\begin{aligned} E[\tilde{r}^4_{1,t}] &\le \bar{c}\left(\sum_{j=1}^{\infty}\sum_{\ell=0}^{K-2}|\tilde{G}_{1,\ell}|\right)^4 < \infty, \\ E[\tilde{r}^4_{2,t}] &\le \bar{c}\left(\sum_{j=1}^{K-2}\sum_{\ell=0}^{K-j-2}|\tilde{G}_{2,\ell}|\right)^4 < \infty \end{aligned}$$

uniformly over K by Lemma A.1(iii) and (iv). The second statement of (i) for $\tilde{r}_t^+$ is proved in exactly the same manner.

(ii) For $i = 1$, we first show that $E[r_{1,t}] = o(1/K^2)$. From the definition of $r_{1,t}$, we have

$$E[|r_{1,t}|] \le \sigma_\xi^2 \sum_{j=-\infty}^{\infty}|\phi_j||\phi_{j-K}|. \tag{36}$$

Noting that

$$\begin{aligned} \sum_{K=-\infty}^{\infty}|K|^2\sum_{j=-\infty}^{\infty}|\phi_j||\phi_{j-K}| &\le 2\sum_{K=-\infty}^{\infty}\sum_{j=-\infty}^{\infty}(|j-K|^2+|j|^2)|\phi_j||\phi_{j-K}| \\ &\le 4\sum_{K=-\infty}^{\infty}|\phi_K|\sum_{j=-\infty}^{\infty}|j|^2|\phi_j| < \infty \end{aligned}$$

because of the 2-summability of $\{\phi_j\}$, we can see that $|K|^2\sum_{j=-\infty}^{\infty}|\phi_j||\phi_{j-K}|$ is a convergence sequence over K. In other words, $K^2\sum_{j=-\infty}^{\infty}|\phi_j||\phi_{j-K}|$ is $o(1)$ as $K \to \infty$ and then from (36), $E[|r_{1,t}|] = o(1/K^2)$ uniformly over t.

Using this result, since

$$\sup_{0\le r\le 1}\left|\frac{1}{\sqrt{T}}\sum_{t=1}^{[Tr]} r_{1,t}\right| \le \frac{1}{\sqrt{T}}\sum_{t=1}^{T}|r_{1,t}|,$$

we obtain

$$E\left[\sup_{0\le r\le 1}\left|\frac{1}{\sqrt{T}}\sum_{t=1}^{[Tr]} r_{1,t}\right|\right] \le \frac{1}{\sqrt{T}}\sum_{t=1}^{T} E\left[|r_{1,t}|\right] = o\left(\frac{\sqrt{T}}{K^2}\right) = o(1).$$

For $i=2$, by Cauchy-Schwarz inequality, we have

$$\begin{aligned}
E[|r_{2,t}|] &\le \left\{E\left[\left(\sum_{|j|\ge K}^{\infty}\phi_j\xi_{t-j}\right)^2\right] E\left[\left(\sum_{\ell=-\infty}^{\infty}\phi_\ell\xi_{t-K-\ell}\right)^2\right]\right\}^{1/2} \\
&= \left\{\sigma_\xi^4\sum_{|j|\ge K}^{\infty}\phi_j^2\sum_{\ell=-\infty}^{\infty}\phi_\ell^2\right\}^{1/2} \\
&= o\left(\frac{1}{K^2}\right)
\end{aligned}$$

by Lemma A.1(i). Then, we have $E[\sup_r |T^{-1/2}\sum_{t=1}^{[Tr]} r_{2,t}|] = o(1)$ in exactly the same manner as the proof for $i=1$.

The case with $i=3$ is shown similarly and we omit the proof.■

From Lemma A.2, the rest we have to show is that the FCLT holds for the first term on the right-hand side of (35). From Theorem 27.14 of Davidson (1994), it is sufficient to show that

$$\frac{\sum_{t=1}^{T} m_t^2}{\sum_{t=1}^{T} E[m_t^2]} \xrightarrow{p} 1, \tag{37}$$

$$\frac{\max_{1\le t\le T}|m_t|}{(\sum_{t=1}^{T} E[m_t^2])^{1/2}} \xrightarrow{p} 0, \tag{38}$$

$$\lim_{T\to\infty}\frac{\sum_{t=1}^{[Tr]} E[m_t^2]}{\sum_{t=1}^{T} E[m_t^2]} \to r \quad \forall\, 0\le r\le 1, \tag{39}$$

where $m_t = \sum_{j=1}^{\infty} G_j\xi_t\xi_{t-j}$.

The condition (37) holds if we show that $T^{-1}\sum_{t=1}^{T}(m_t^2 - E[m_t^2])) \xrightarrow{p} 0$, which is proved using

Chebyshev inequality by showing that

$$
\begin{aligned}
E\left[\left\{\frac{1}{T}\sum_{t=1}^{T}(m_t^2 - E[m_t^2])\right\}^2\right] &= \frac{1}{T^2}\sum_{t=1}^{T} E\left[(m_t^2 - E[m_t^2])^2\right] \\
&\quad + \frac{2}{T^2}\sum_{s=1}^{T-1}\sum_{t=s+1}^{T} E\left[(m_t^2 - E[m_t^2])(m_{t-s}^2 - E[m_{t-s}^2])\right] \\
&\to 0. \qquad (40)
\end{aligned}
$$

For the first term on the right-hand side of (40),

$$
\begin{aligned}
\frac{1}{T^2}\sum_{t=1}^{T} E\left[\left(m_t^2 - E[m_t^2]\right)^2\right] &= \frac{1}{T^2}\sum_{t=1}^{T} E\left[\left\{\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} G_i G_j \left(\xi_t^2 \xi_{t-i}\xi_{t-j} - \sigma_\xi^2 E[\xi_{t-i}\xi_{t-j}]\right)\right\}^2\right] \\
&\leq \frac{\bar{c}}{T}\left(\sum_{j=1}^{\infty}|G_j|\right)^4 \to 0. \qquad (41)
\end{aligned}
$$

For the second term, note that for $s > 0$,

$$
\begin{aligned}
&E[(m_t^2 - E[m_t^2])(m_{t-s}^2 - E[m_{t-s}^2])] \\
&= \sum_{i_1=1}^{\infty}\sum_{j_1=1}^{\infty}\sum_{i_2=1}^{\infty}\sum_{j_2=1}^{\infty} G_{i_1}G_{i_2}G_{i_3}G_{i_4} E\left[(\xi_t^2\xi_{t-i_1}\xi_{t-j_1} - \sigma_\xi^2 E[\xi_{t-i_1}\xi_{t-j_1}])\right. \\
&\qquad\qquad \left.(\xi_{t-s}^2\xi_{t-s-i_2}\xi_{t-s-j_2} - \sigma_\xi^2 E[\xi_{t-s-i_2}\xi_{t-s-j_2]}])\right].
\end{aligned}
$$

The expectation becomes

$$
\begin{aligned}
&E\left[(\xi_t^2\xi_{t-i_1}\xi_{t-j_1} - \sigma_\xi^2 E[\xi_{t-i_1}\xi_{t-j_1}])(\xi_{t-s}^2\xi_{t-s-i_2}\xi_{t-s-j_2} - \sigma_\xi^2 E[\xi_{t-s-i_2}\xi_{t-s-j_2]}])\right] \\
&= E\left[\sigma_\xi^2(\xi_{t-i_1}\xi_{t-j_1} - E[\xi_{t-i_1}\xi_{t-j_1}])(\xi_{t-s}^2 - \sigma_\xi^2)\xi_{t-s-i_2}\xi_{t-s-j_2}\right] \\
&\quad + E\left[\sigma_\xi^2(\xi_{t-i_1}\xi_{t-j_1} - E[\xi_{t-i_1}\xi_{t-j_1}])\sigma_\xi^2(\xi_{t-s-i_2}\xi_{t-s-j_2} - E[\xi_{t-s-i_2}\xi_{t-s-j_2]}])\right].
\end{aligned}
$$

Since $\{\xi_t\}$ is an independent sequence, the first expectation takes non-zero values when i) $i_1 = j_1 = s$ and $i_2 = j_2$, ii) $i_1 = s + i_2$ and $j_1 = s + j_2$, (iii) $i_1 = s + j_2$ and $j_1 = s + j_2$, while for the second expectation, it is sufficient to consider either iv) $i_1 = s + i_2$ and $j_1 = s + j_2$ or (v) $i_1 = s + j_2$ and $j_1 = s + j_2$. Therefore, we can see that

$$
\left|E[(m_t^2 - E[m_t^2])(m_{t-s}^2 - E[m_{t-s}^2])]\right| \leq \bar{c}\left[G_s^2\sum_{j_2=1}^{\infty} G_{j_2}^2 + \left(\sum_{i_2=1}^{\infty}|G_{s+i_2}||G_{i_2}|\right)^2\right],
$$

and thus,

$$\left| \frac{1}{T^2} \sum_{s=1}^{T-1} \sum_{t=s+1}^{T} E\left[(m_t^2 - E[m_t^2])(m_{t-s}^2 - E[m_{t-s}])^2 \right] \right| \tag{42}$$

$$\leq \quad \frac{\bar{c}}{T} \left[\sum_{s=1}^{T-1} G_s^2 \sum_{j_2=1}^{\infty} G_{j_2}^2 + \sum_{s=1}^{T-1} \left(\sum_{i_2=1}^{\infty} |G_{s+i_2}||G_{i_2}| \right)^2 \right] \leq \frac{\bar{c}}{T} \left[\left(\sum_{j_2=1}^{\infty} G_{j_2}^2 \right)^2 + \left(\sum_{i_2=1}^{\infty} |G_{i_2}| \right)^4 \right] \to 0.$$

Then, (40) holds from (41) and (42).

To prove (38), we note that $E[m_t^2] = \sigma_\xi^4 \sum_{j=1}^{\infty} G_j^2 < \infty$ and then the denominator of (38) is $O(\sqrt{T})$. On the other hand,

$$\begin{aligned} P\left(\max_{1 \leq t \leq T} \frac{1}{\sqrt{T}} |m_t| \geq \varepsilon \right) \quad &\leq \quad T\, P\left(\frac{1}{\sqrt{T}} |m_t| \geq \varepsilon \right) \\ &\leq \quad \frac{1}{\varepsilon^4 T} E[m_t^4] = O\left(\frac{1}{T} \right), \end{aligned}$$

because $E[m_t^4]$ is bounded uniformly in t, T and M. Therefore, we obtain (38).

Finally, we can see that (39) holds even in finite samples because of stationarity of m_t.■

Proof of Lemma 3

Let $\tilde{D}_T = diag\{\sqrt{T}, T\, I_{p_x}\}$ (constant case) or $\tilde{D}_T = \{\sqrt{T}, T\sqrt{T}, I_{p_x}\}$ (trend case), $\|B\| = [tr(B'B)]^{1/2}$ and $\|B\|_1 = \sup\{\|Bx\| : \|x\| \leq 1\}$ for a matrix B. We will show that only $R_{\beta,T}$ yields the non-zero bias whereas $R_{\Pi,T}$ and R_T are negligible, using the following lemma:

Lemma A.3 *Suppose that Assumptions 1, 2 and 3 hold. Under the null hypothesis, as $T \to \infty$, (i) $\tilde{D}_T^{-1}(\hat{\beta}-\beta) \xrightarrow{d} \left(\int_0^1 \tilde{B}(r)\tilde{B}'(r)dr\right)^{-1} \int_0^1 \tilde{B}(r)dB_\eta(r)$, (ii) $\|\hat{\Pi}-\Pi\|^2 = O_p(M/T)$, (iii) $\tilde{D}^{-1}\sum_{t=K+1}^{T} X_{t-K}\eta_t \xrightarrow{d} \int_0^1 \tilde{B}(r)dB_\eta(r)$ and $\tilde{D}^{-1}\sum_{t=K+1}^{T} X_t\eta_{t-K} \xrightarrow{d} \int_0^1 \tilde{B}(r)dB_\eta(r)$, (iv) $\|T^{-1/2}\sum_{t=K+1}^{T} V_{t-K}\eta_t\| = \|T^{-1/2}\sum_{t=K+1}^{T} V_t\eta_{t-K}\|$ $O_p(M^{1/2})$, (v) $\|\sum_{t=K+1}^{T} \eta_t e_{t-K}\| = \|\sum_{t=K+1}^{T} \eta_{t-K} e_t\| = o_p(1)$, (vi) $\tilde{D}^{-1}\sum_{t=K+1}^{T} X_t X_{t-K}'\tilde{D}^{-1} \xrightarrow{d} \int_0^1 \tilde{B}(r)\tilde{B}'(r)dr$,*
(vii) $\|\tilde{D}_T^{-1}\sum_{t=K+1}^{T} X_t V_{t-K}'\| = \|\tilde{D}_T\sum_{t=K+1}^{T} X_{t-K} V_t'\| = O_p(M^{1/2})$, (viii) $\|\tilde{D}_T\sum_{t=K+1}^{T} X_t e_{t-K}\| = \|\tilde{D}_T\sum_{t=K+1}^{T} X_{t-K} e_t\| = o_p(1)$, (ix) $\|T^{-1/2}\sum_{t=K+1}^{T} V_t V_{t-K}'\| = O_p(M)$, (x) $\|\left(T^{-1}\sum_{t=K+1}^{T} V_t V_t'\right)^{-1} - \Gamma_x^{-1}\|_1 = O_p(M/\sqrt{T})$, (xi) $\|\sum_{t=K+1}^{T} V_t e_{t-K}\| = \|\sum_{t=K+1}^{T} V_{t-K} e_t\| = o_p(M^{1/2})$, (xii) $\|\sum_{t=K+1}^{T} e_t e_{t-K}\| = o_p(1)$, where $\tilde{B}(r) = [1, B'(r)]'$ (constant case) or $\tilde{B}(r) = [1, r, B'(r)]'$ (trend case) with $B(r)$ being a p_x-dimensional Brownian motion with the variance given by $\lim_{T\to\infty} E[T^{-1/2}x_T]$, $B_\eta(r)$ is a one-dimensional Brownian motion independent of $B(r)$ with the variance given by $\omega_\eta^2 = \lim_{T\to\infty} E[(T^{-1/2}\sum_{t=1}^{T}\eta_t)^2]$ and $\Gamma_x = E[V_t V_t']$.

Proof of Lemma A.3: All the results, except for (v), (ix) and (xi), are obtained by Saikkonen (1991) using the FCLT with K going to infinity slower than T. For (v), we can see that

$$\left| \sum_{t=K+1}^{T} \eta_t e_{t-K} \right| \leq \sup_{1\leq t\leq T} |e_t| \sum_{t=1}^{T} |\eta_t|.$$

Note that $\sum_{t=1}^{T} |\eta_t| = O_p(T)$ while

$$\begin{aligned} P\left(\sup_{1\leq t\leq T} |e_t| \geq \varepsilon \right) &\leq TP\left(|e_t| \geq \varepsilon\right) \\ &\leq \frac{T}{\varepsilon^4} E[e_t^4] \\ &\leq \frac{\bar{c}T}{\varepsilon^4} \left(\sum_{|j|\geq K} \|\pi_j\| \right)^4 \\ &= \frac{\bar{c}T}{\varepsilon^4} o\left(\frac{1}{T^2}\right) = o\left(\frac{1}{T}\right), \end{aligned} \tag{43}$$

where the second last equality is obtained by (8).We thus obtain (v).

For (ix), each block element is expressed as $T^{-1/2} \sum_{t=K+1}^{T} v_{t-i} v'_{t-K-j}$ for $i, j = -M, \cdots, M$. Since $(t-i) - (t-K-j) = K - i + j \geq K - 2M$, we can see that the time difference diverges to infinity at a rate of K because $M/K \to 0$ by Assumptions 2 and 3. Because the conditions for the FCLT given by HML (2003) are satisfied, we can see that each element is $O_p(1)$, which implies (ix).

(xi) is proved by noting that

$$E\left[\left\| \sum_{t=K+1}^{T} V_t e_{t-K} \right\| \right] \leq \sup |e_t| \sum_{t=1}^{T} E[\|V_t\|] = o_p(\sqrt{M}),$$

because $\sup_t |e_t| = o_p(1/T)$ by (43).■

We first evaluate $R_{\beta,T}$. Using Lemma A.3(i), (iii) and (vi), we have

$$R_{\beta,T} \xrightarrow{d} -\left(\int_0^1 \tilde{B}(r) dB_\eta(r) \right)' \left(\int_0^1 \tilde{B}(r)\tilde{B}'(r)dr \right)^{-1} \left(\int_0^1 \tilde{B}(r) dB_\eta(r) \right). \tag{44}$$

Since $\int_0^1 \tilde{B}(r)dB_\eta(r)|\tilde{B}(\cdot) \sim N\left(0, \omega_\eta^2 \int_0^1 \tilde{B}(r)\tilde{B}'(r)dr\right)$, we can see that the right-hand side of (44) is distributionally equal to $-\omega_\eta^2$ times a chi-square distribution with $(p_c + p_x)$ degrees of freedom. As a result, $E[R_{\beta,T}]$ can be approximated by $-\omega_\eta^2(p_c + p_x)$.

For $R_{\Pi,T}$, the first term becomes

$$\begin{aligned} \left\| (\hat{\Pi} - \Pi)' \sum_{t=K+1}^{T} V_t V_{t-K} (\hat{\Pi} - \Pi) \right\| &\leq \left\| \hat{\Pi} - \Pi \right\|^2 \left\| \sum_{t=K+1}^{T} V_t V_{t-K} \right\| \\ &= O_p\left(\frac{M^2}{\sqrt{T}} \right) = o_p(1), \end{aligned}$$

using Lemma A.3 (ii) and (ix) and Assumption 2.

For the second term of $R_{\Pi,T}$, since it can be shown that

$$\left\| \sqrt{T}(\hat{\Pi} - \Pi) - \left(\frac{1}{T} \sum_{t=1}^{T} V_t V_t' \right)^{-1} \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} V_t \eta_t \right) \right\| \leq O_p \left(\sqrt{\frac{M}{T}} \right),$$

while

$$\left\| \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \eta_t V_t' \right) \left[\left(\frac{1}{T} \sum_{t=1}^{T} V_t V_t' \right)^{-1} - \Gamma_x^{-1} \right] \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} V_{t-K} \eta_t \right) \right\| = O_p \left(\frac{M^2}{\sqrt{T}} \right) = o_p(1)$$

by Lemma A.3 (iv) and (x), it is sufficient to evaluate

$$\begin{aligned} & \left| E \left[\left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \eta_t V_t' \right) \Gamma_x^{-1} \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} V_{t-K} \eta_t \right) \right] \right| \\ \leq & \quad \sup |\Gamma_x^{-1}(i,j)| \frac{1}{T} \sum_{t=1}^{T} \sum_{\ell=t-T}^{t-1} \left| E \left[V_t' V_{t-K-\ell} \eta_t \eta_{t-\ell} \right] \right|. \end{aligned} \quad (45)$$

To evaluate the right-hand side of (45), we express η_t using ε_t such that

$$\eta_t = \sum_{j=-\infty}^{\infty} \psi_j' \varepsilon_{t-j}, \quad \text{where} \quad \sum_{j=-\infty}^{\infty} |j|^2 \|\psi_j\| \leq \infty$$

with $\{\psi_j\}_{j=-\infty}^{\infty}$ is a sequence of $p_x + 1$-dimensional coefficient vectors, because

$$\eta_t = u_t - \sum_{j=-\infty}^{\infty} \pi_j' v_{t-j} \quad \text{with} \quad u_t = \sum_{j=0}^{\infty} \Phi_{1,j} \varepsilon_{t-j} \quad \text{and} \quad v_t = \sum_{j=0}^{\infty} \Phi_{2,j} \varepsilon_{t-j},$$

where Φ_j is partitioned into $\Phi_j = [\Phi_{1,j}', \Phi_{2,j}']'$. Then, by focusing on the term $v_t' v_{t-K-\ell}$ in $V_t' V_{t-K-\ell}$, we can see that

$$\begin{aligned} \tilde{R}_{\Pi,\ell} & \equiv E \left[v_t' v_{t-K-\ell} \eta_t \eta_{t-\ell} \right] \\ & = E \left[\left(\sum_{j_1=0}^{\infty} \Phi_{2,j_1} \varepsilon_{t-j_1} \right)' \left(\sum_{j_2=0}^{\infty} \Phi_{2,j_2} \varepsilon_{t-K-\ell-j_2} \right) \left(\sum_{i_1=-\infty}^{\infty} \psi_{i_1}' \varepsilon_{t-i_1} \right) \left(\sum_{i_2=-\infty}^{\infty} \psi_{i_2}' \varepsilon_{t-\ell-i_2} \right) \right]. \end{aligned}$$

We note that the expectation takes non-zero values when (i) $j_1 = K + \ell + j_2$, $i_1 = \ell + i_2$ and $i_2 \neq K + j_2$, (ii) $i_1 = j_1$, $i_2 = K + j_2$ and $j_1 \neq K + \ell + j_2$, (iii) $i_1 = K + \ell + j_2$, $i_2 = j_1 - \ell$ and $j_1 \neq K + \ell + j_2$, and (iv) $i_1 = K + \ell + j_2$, $i_2 = K + j_2$ and $j_1 = K + \ell + j_2$.

In case (i), for $\ell \geq 0$, the sum of $\tilde{R}_{\Pi,\ell}$ becomes

$$
\begin{aligned}
\left|\sum_{\ell=0}^{\infty} \tilde{R}_{\Pi,\ell}\right| &\leq \bar{c} \sum_{\ell=0}^{\infty} \sum_{j_2=0}^{\infty} \|\Phi_{2,K+\ell+j_2}\| \|\Phi_{j_2}\| \sum_{i_2=-\infty}^{\infty} \|\psi_{\ell+i_2}\| \|\psi_{i_2}\| \\
&\leq \bar{c} \sum_{\ell=0}^{\infty} \sum_{j_2=0}^{\infty} \|\Phi_{2,K+\ell+j_2}\| \left(\sum_{j=0}^{\infty} \|\Phi_{2,j}\|\right) \left(\sum_{i_2=-\infty}^{\infty} \|\psi_{i_2}\|\right)^2 \\
&\leq \bar{c} \sum_{j_2=K}^{\infty} (j_2 - K + 1) \|\Phi_{2,j_2}\| = o\left(\frac{1}{K}\right), \qquad (46)
\end{aligned}
$$

because $\{\Phi_j\}$ is 2-summable.

On the other hand, for $\ell = -1, -2, \cdots, -K$, we have

$$
\begin{aligned}
\left|\sum_{\ell=-K}^{-1} \tilde{R}_{\Pi,\ell}\right| &\leq \bar{c} \sum_{\ell=-K}^{[-K/2]} \sum_{j_2=0}^{\infty} \|\Phi_{2,K+\ell+j_2}\| \|\Phi_{j_2}\| \sum_{i_2=-\infty}^{\infty} \|\psi_{\ell+i_2}\| \|\psi_{i_2}\| \\
&\quad + \bar{c} \sum_{\ell=[-K/2]+1}^{-1} \sum_{j_2=0}^{\infty} \|\Phi_{2,K+\ell+j_2}\| \|\Phi_{j_2}\| \sum_{i_2=-\infty}^{\infty} \|\psi_{\ell+i_2}\| \|\psi_{i_2}\| \\
&\leq \bar{c} \left(\sum_{j_2=0}^{\infty} \|\Phi_{j_2}\|\right)^2 \sum_{\ell=-K}^{[-K/2]} \sum_{i_2=-\infty}^{\infty} \|\psi_{\ell+i_2}\| \|\psi_{i_2}\| \\
&\quad + \bar{c} \left(\sum_{\ell=[-K/2]+1}^{-1} \sum_{j_2=0}^{\infty} \|\Phi_{2,K+\ell+j_2}\|\right) \left(\sum_{j_2=0}^{\infty} \|\Phi_{j_2}\|\right) \left(\sum_{i_2=-\infty}^{\infty} \|\psi_{i_2}\|\right)^2 \\
&= o\left(\frac{1}{K}\right) + o\left(\frac{1}{K}\right), \qquad (47)
\end{aligned}
$$

where the last relation holds because

$$
\begin{aligned}
\sum_{K=-\infty}^{\infty} |K|^2 \sum_{i_2=-\infty}^{\infty} \|\psi_{i_2-K}\| \|\psi_{i_2}\| &\leq 2 \sum_{K=-\infty}^{\infty} \sum_{i_2=-\infty}^{\infty} (|i_2|^2 + |i_2 - K|^2) \|\psi_{i_2-K}\| \|\psi_{i_2}\| \\
&\leq 4 \sum_{K=-\infty}^{\infty} \|\psi_K\| \sum_{i_2=-\infty}^{\infty} |i_2|^2 \|\psi_{i_2}\| < \infty,
\end{aligned}
$$

which implies $|K|^2 \sum_{i_2=-\infty}^{\infty} \|\psi_{i_2-K}\| \|\psi_{i_2}\| = o(1)$ or, equivalently, $\sum_{i_2=-\infty}^{\infty} \|\psi_{i_2-K}\| \|\psi_{i_2}\| = o(1/K^2)$, while

$$
\sum_{\ell=[-K/2]+1}^{-1} \sum_{j_2=0}^{\infty} \|\Phi_{2,K+\ell+j_2}\| \leq \left[\frac{K}{2}\right] \sum_{j_2=[K/2]}^{\infty} \|\Phi_{2,j_2}\| = o\left(\frac{1}{K}\right)
$$

because of 2-summability of $\{\Phi_{2,j}\}$.

For $\ell \le -K-1$,

$$\begin{aligned}
\left|\sum_{\ell=-\infty}^{-K+1} \tilde{R}_{\Pi,\ell}\right| &\le \bar{c} \sum_{\ell=-\infty}^{-K-1} \sum_{j_1=0}^{\infty} \|\Phi_{2,j_1}\| \|\Phi_{j_1-K-\ell}\| \sum_{i_2=-\infty}^{\infty} \|\psi_{\ell+i_2}\| \|\psi_{i_2}\| \\
&\le \bar{c} \sum_{\ell=-\infty}^{-K-1} \sum_{j_1=0}^{\infty} \|\Phi_{j_1-K-\ell}\| \left(\sum_{j_1=0}^{\infty} \|\Phi_{2,j_1}\|\right) \left(\sum_{i_2=-\infty}^{\infty} \|\psi_{i_2}\|\right)^2 \\
&\le \bar{c} \sum_{j_1=K}^{\infty} j_1 \|\Phi_{j_1}\| = o\left(\frac{1}{K}\right). \qquad (48)
\end{aligned}$$

From (46)–(48), we have $\left|\sum_{\ell=-\infty}^{\infty} \tilde{R}_{\Pi,\ell}\right| = o(1/K)$ in case (i).

In case (ii), we first note that

$$E[v_s \eta_t] = \sum_{j=0}^{\infty} \Phi_{2,j} \Sigma_\varepsilon \psi_{j+\ell} = 0 \quad \forall \ell = 0, \pm 1, \pm 2, \cdots. \qquad (49)$$

Then, we have for $\ell \ge 0$,

$$\begin{aligned}
\left|\sum_{\ell=0}^{\infty} \tilde{R}_{\Pi,\ell}\right| &= \left|\sum_{\ell=0}^{\infty} \sum_{\substack{j_2=0 \\ j_1 \ne K+\ell+j_2}}^{\infty} \sum_{j_1=0}^{\infty} \psi_{j_1}' \Sigma_\varepsilon \Phi_{2,j_1}' \Phi_{2,j_2} \Sigma_\varepsilon \psi_{K+j_2}\right| \\
&= \left|\sum_{\ell=0}^{\infty} \sum_{j_2=0}^{\infty} \psi_{K+\ell+j_2}' \Sigma_\varepsilon \Phi_{2,K+\ell+j_2}' \Phi_{2,j_2} \Sigma_\varepsilon \psi_{K+j_2}\right| \\
&\le \bar{c} \sum_{\ell=0}^{\infty} \sum_{j_2=0}^{\infty} \|\psi_{K+\ell+j_2}\| \|\Phi_{2,K+\ell+j_2}\| \|\Phi_{2,j_2}\| \psi_{K+j_2}\| \\
&\le \bar{c} \left(\sum_{j_2=0}^{\infty} j_2 \|\psi_{K+j_2}\|\right) \left(\sum_{j_2=0}^{\infty} j_2 \|\Phi_{2,K+j_2}\|\right) \left(\sum_{j_2=0}^{\infty} \|\Phi_{2,j_2}\|\right) \left(\sum_{j_2=0}^{\infty} \|\psi_{K+j_2}\|\right) \\
&= o\left(\frac{1}{K^2}\right),
\end{aligned}$$

where the second equality holds using (49).

Similarly for $\ell = -1, \cdots, -K$,

$$\begin{aligned}
\left|\sum_{\ell=-K}^{-1} R_{\Pi,t}\right| &\le \bar{c} \sum_{\ell=-K}^{-1} \sum_{j_2=0}^{\infty} \|\psi_{K+\ell+j_2}\| \|\Phi_{2,K+\ell+j_2}\| \|\Phi_{2,j_2}\| \psi_{K+j_2}\| \\
&\le \bar{c} \left(\sum_{\ell=-K}^{-1} \sum_{j_2=0}^{\infty} \|\Phi_{2,K+\ell+j_2}\|\right) \left(\sum_{j_2=-\infty}^{\infty} \|\psi_{j_2}\|\right) \left(\sum_{j_2=0}^{\infty} \|\Phi_{2,j_2}\|\right) \left(\sum_{j_2=0}^{\infty} \|\psi_{K+j_2}\|\right) \\
&\le \bar{c} \left(\sum_{j_2=0}^{K-1} (j_2+1) \|\Phi_{2,j_2}\| + K \sum_{j_2=K}^{\infty} \|\Phi_{2,j_2}\|\right) \left(\sum_{j_2=0}^{\infty} \|\psi_{K+j_2}\|\right) = o\left(\frac{1}{K}\right),
\end{aligned}$$

while for $\ell \leq -K-1$,

$$
\begin{aligned}
\left|\sum_{\ell=-\infty}^{-K-1} R_{\Pi,t}\right| &= \left|\sum_{\ell=-\infty}^{-K-1} \sum_{\substack{j_2=0 \\ j_2 \neq j_1-K-\ell}}^{\infty} \sum_{j_1=0}^{\infty} \psi_{j_1}' \Sigma_\varepsilon \Phi_{2,j_1}' \Phi_{2,j_2} \Sigma_\varepsilon \psi_{K+j_2}\right| \\
&= \left|\sum_{\ell=-\infty}^{-K-1} \sum_{j_1=0}^{\infty} \psi_{j_1}' \Sigma_\varepsilon \Phi_{2,j_1}' \Phi_{2,j_1-K-\ell} \Sigma_\varepsilon \psi_{j_1-\ell}\right|
\end{aligned}
$$

$$
\begin{aligned}
&\leq \bar{c}\left(\sum_{j_1=0}^{\infty} \|\psi_{j_1}\|\right)\left(\sum_{j_1=0}^{\infty} \|\Phi_{2,j_1}\|\right)\left(\sum_{j_1=0}^{\infty} j_1 \|\Phi_{2,j_1}\|\right)\left(\sum_{j_1=K+1}^{\infty} (j_1-K)\|\psi_{j_1}\|\right) \\
&= o\left(\frac{1}{K}\right).
\end{aligned}
$$

We then have $\left|\sum_{\ell=-\infty}^{\infty} \tilde{R}_{\Pi,\ell}\right| = o(1/K)$ in case (ii).

In exactly the same way, we have the same order in cases (iii) and (iv), so that $\left|\sum_{\ell=-\infty}^{\infty} \tilde{R}_{\Pi,\ell}\right| = o(1/K)$ in general. Then, we can see that the right-hand side of (45) is $o(M/K) = o(1)$ by Assumptions 2 and 3, so that the second term of $R_{\Pi,T}$ is $o_p(1)$. Similarly, we can show that the third term of $R_{\Pi,T}$ is $o_p(1)$.

Using Lemma A.3, it is not difficult to see that $R_T = o_p(1)$. As a result, we obtain the bias.$\blacksquare$

Proof of Theorem 2

For model (16), the regression residuals $\hat{\eta}_t^*$ can be expressed as

$$
\hat{\eta}_t^* = \dot{\eta}_t + e_t - (\hat{\beta}-\beta)' X_t - (\hat{\Pi}-\Pi)' V_t.
$$

In exactly the same manner as the proof of Theorem 1 of Kurozumi and Hayakawa (2009), it can be shown that

$$
\frac{1}{T^{\vartheta+1}} \sum_{t=K+j+1}^{T} \hat{\eta}_{t-j}^* \hat{\eta}_{t-j-K}^* = \frac{1}{T^{\vartheta+1}} \sum_{t=K+j+1}^{T} \dot{\eta}_{t-j} \dot{\eta}_{t-j-K} + o_p(1)
$$

for $j = 0, 1, 2, \cdots$ and thus we proceed with the proof using $\dot{\eta}_t$ instead of $\hat{\eta}_t^*$ in the following.

We first evaluate the numerator of $\hat{S}_K$. Since

$$
\dot{\eta}_t = \dot{\rho}^K \dot{\eta}_{t-K} + q_{K,t}, \tag{50}
$$

where $q_{K,t} = \sum_{\ell=t-K+1}^{t} \dot{\rho}^{t-\ell} \xi_\ell$ from the definition, we can see that

$$
\sum_{t=K+1}^{T} \dot{\eta}_t \dot{\eta}_{t-K} = \dot{\rho}^K \sum_{t=K+1}^{T} \dot{\eta}_{t-K}^2 + \sum_{t=K+1}^{T} q_{K,t} \dot{\eta}_{t-K}. \tag{51}
$$

Lemma 2(a) in Kurozumi and Hayakawa (2009) proved that

$$\frac{1}{T^{1+\vartheta}} \sum_{t=K+1}^{T} \dot{\eta}_{t-K}^2 \xrightarrow{p} \frac{\sigma_\xi^2}{2c_1} \tag{52}$$

while

$$\dot{\rho}^K = \left(1 - \frac{c_1}{T^\vartheta}\right)^K \to \begin{cases} 0 & : \text{ if } \frac{K}{T^\vartheta} \to \infty \\ e^{-c_1\tau} & : \text{ if } \frac{K}{T^\vartheta} \to \tau. \end{cases} \tag{53}$$

Then, if we show that the second term on the right hand side of (51) is $o_p(T^{\vartheta+1})$, we have, from (52) and (53)

$$\frac{1}{T^{\vartheta+1}} \sum_{t=K+1}^{T} \dot{\eta}_t \dot{\eta}_{t-K} \xrightarrow{p} \begin{cases} 0 & : \text{ if } \frac{K}{T^\vartheta} \to \infty \\ \frac{\sigma_\xi^2}{2c_1} e^{-c_1\tau} & : \text{ if } \frac{K}{T^\vartheta} \to \tau. \end{cases} \tag{54}$$

To derive the order of the second term of (51), we evaluate

$$E\left[\left\{\sum_{t=K+1}^{T} q_{K,t}\dot{\eta}_{t-K}\right\}^2\right] = \sum_{t=K+1}^{T} E\left[q_{K,t}^2 \dot{\eta}_{t-K}^2\right] + 2\sum_{s=K+1}^{T-1}\sum_{t=s+1}^{T} E\left[q_{K,s}q_{K,t}\dot{\eta}_{t-K}\dot{\eta}_{s-K}\right]. \tag{55}$$

Since $q_{K,t}$ is independent of $\dot{\eta}_{t-K}$, the first term on the right hand side of (55) becomes

$$\begin{aligned} \sum_{t=K+1}^{T} E\left[q_{K,t}^2 \dot{\eta}_{t-K}^2\right] &= \sum_{t=K+1}^{T} E[q_{K,t}^2] E[\dot{\eta}_{t-K}^2] \\ &= \sigma_\xi^4 \sum_{t=K+1}^{T} \frac{1-\dot{\rho}^{2K}}{1-\dot{\rho}^2} \frac{1-\dot{\rho}^{2(t-K)}}{1-\dot{\rho}^2} \\ &= O(K \wedge T^\vartheta) \times O(T^{\vartheta+1}) \end{aligned} \tag{56}$$

where we used the fact that

$$\frac{1-\dot{\rho}^{2K}}{1-\dot{\rho}^2} = \begin{cases} O(T^\vartheta) & : \text{ if } \frac{K}{T^\vartheta} \to \infty \\ O(K) & : \text{ if } \frac{K}{T^\vartheta} \to \tau. \end{cases} \tag{57}$$

Similarly, for the second term on the right hand side of (55), it can be shown by complicated but direct calculations that

$$\sum_{s=K+1}^{T-1}\sum_{t=s+1}^{T} E\left[q_{K,s}q_{K,t}\dot{\eta}_{t-K}\dot{\eta}_{s-K}\right] = O(K \wedge T^\vartheta) \times O(T).$$

As a result, we have

$$E\left[\left\{\sum_{t=K+1}^{T} q_{K,t}\dot{\eta}_{t-K}\right\}^2\right] = O(K \wedge T^\vartheta) \times O(T^{\vartheta+1})$$

and thus the second term on the right hand side of (51) is $o_p(T^{\vartheta+1})$.

To evaluate the denominator of the test statistic, we first note that, using $\dot{\eta}_t$ instead of $\hat{\eta}_t^*$, the long-run variance estimator is expressed as

$$\frac{1}{T}\sum_{t=K+1}^{T} \dot{\eta}_t^2 \dot{\eta}_{t-K}^2 + 2\sum_{j=1}^{J}\left(1 - \frac{j}{J+1}\right)\frac{1}{T}\sum_{t=K+j+1}^{T} \dot{\eta}_t \dot{\eta}_{t-K} \dot{\eta}_{t-j} \dot{\eta}_{t-K-j}.$$

Using (50), the expectation of $\dot{\eta}_t^2\dot{\eta}_{t-K}^2$ can be expressed as

$$
\begin{aligned}
\frac{1}{T}\sum_{t=K+1}^{T} E\left[\dot{\eta}_t^2\dot{\eta}_{t-K}^2\right] &= \frac{1}{T}\sum_{t=K+1}^{T} E\left[\left(\dot{\rho}^K\dot{\eta}_{t-K}+q_{K,t}\right)^2\dot{\eta}_{t-K}^2\right] \\
&= \dot{\rho}^{2K}\frac{1}{T}\sum_{t=K+1}^{T} E[\dot{\eta}_{t-K}^4] + \frac{1}{T}\sum_{t=K+1}^{T} E[\dot{\eta}_{t-K}^2]E[q_{K,t}^2]. \qquad (58)
\end{aligned}
$$

Since $E[\xi_t^3]=0$, the fourth moment of $\dot{\eta}_t$ becomes

$$
\begin{aligned}
E[\dot{\eta}_t{}^4] &= E\left[\left(\sum_{\ell=0}^{t-1}\dot{\rho}^\ell\xi_{t-\ell}\right)^4\right] \\
&= E[\xi_t^4]\sum_{\ell=0}^{t-1}\dot{\rho}^{4\ell} + 3\left[2\sigma_\xi^4\sum_{\ell=0}^{t-2}\dot{\rho}^{2\ell}\sum_{m=\ell+1}^{t-1}\dot{\rho}^{2m}\right] \\
&= E[\xi_t^4]\frac{1-\dot{\rho}^{4t}}{1-\dot{\rho}^4} + \frac{6\sigma_\xi^4(\dot{\rho}^2-\dot{\rho}^{4t-2})}{(1-\dot{\rho}^2)(1-\dot{\rho}^4)} - \frac{6\sigma_\xi^4(\dot{\rho}^{2t}-\dot{\rho}^{4t-2})}{(1-\dot{\rho}^2)^2}.
\end{aligned}
$$

Then, it can be shown that

$$
\dot{\rho}^{2K}\frac{1}{T}\sum_{t=K+1}^{T} E[\dot{\eta}_{t-K}^4] = \frac{3\sigma_\xi^4}{4c_1^2}\dot{\rho}^{2K}\left(T^{2\vartheta}+o(T^{2\vartheta})\right). \qquad (59)
$$

From (53), we can see that (59) is $o(T^{2\vartheta})$ if $K/T^\vartheta\to\infty$ while it is $O(T^{2\vartheta})$ if $K/T^\vartheta\to\tau$.

For the second term of (58), we can see that

$$
\begin{aligned}
\frac{1}{T}\sum_{t=K+1}^{T} E[\dot{\eta}_{t-K}^2]E[q_{K,t}^2] &= \frac{\sigma_\xi^4}{T}\sum_{t=K+1}^{T}\frac{1-\dot{\rho}^{2(t-K)}}{1-\dot{\rho}^2}\frac{1-\dot{\rho}^{2K}}{1-\dot{\rho}^2} \\
&= \frac{\sigma_\xi^4}{2c_1}T^\vartheta\times O(K\wedge T^\vartheta)\left(1+o(1)\right), \qquad (60)
\end{aligned}
$$

where the last equality holds from (57). Then, we can see that (60) is $O(T^{2\vartheta})$ if $K/T^\vartheta\to\infty$ while it is $O(T^{2\vartheta})$ if $K/T^\vartheta\to\tau\neq 0$ or $o(T^{2\vartheta})$ if $K/T^\vartheta\to 0$.

Combining the orders of (59) and (60), we can see that $T^{-(1+2\vartheta)}\sum_{t=K+1}^{T}E[\dot{\eta}_t^2\dot{\eta}_{t-K}^2]$ has a non-zero limit irrespective of the relative order of K to T^ϑ.

Similarly, using (50),

$$\begin{aligned}
&\frac{1}{T}\sum_{t=K+j+1}^{T} E\left[\dot{\eta}_t\dot{\eta}_{t-K}\dot{\eta}_{t-j}\dot{\eta}_{t-K-j}\right] \\
&= \frac{1}{T}\sum_{t=K+j+1}^{T} E\left[\left(\dot{\rho}^j\dot{\eta}_{t-j}+q_{j,t}\right)\left(\dot{\rho}^j\dot{\eta}_{t-K-j}+q_{K+j,t-K}\right)\dot{\eta}_{t-j}\dot{\eta}_{t-K-j}\right] \\
&= \rho^{2j}\frac{1}{T}\sum_{t=K+j+1}^{T} E[\dot{\eta}_{t-j}^2\dot{\eta}_{t-K-j}^2]+\rho^j\frac{1}{T}\sum_{t=K+j+1}^{T} E[q_{K+j,t-K}\dot{\eta}_{t-j}^2\dot{\eta}_{t-K-j}] \\
&\quad +\dot{\rho}^j\frac{1}{T}\sum_{t=K+j+1}^{T} E[q_{j,t}\dot{\eta}_{t-j}\dot{\eta}_{t-K-j}^2]+\frac{1}{T}\sum_{t=K+j+1}^{T} E[q_{j,t}q_{K+j,t-K}\dot{\eta}_{t-j}\dot{\eta}_{t-K-j}]. \qquad (61)
\end{aligned}$$

Note that $T^{-(1+2\vartheta)}\sum_{t=j+K+1}^{T} E[\dot{\eta}_{t-j}^2\dot{\eta}_{t-j-K}^2]$ is $O(1)$ as proved just before while

$$\begin{aligned}
\sum_{j=1}^{J}\left(1-\frac{j}{J+1}\right)\dot{\rho}^{2j} &= \frac{1}{J+1}\frac{J\dot{\rho}^2(1-\dot{\rho}^2)-\dot{\rho}^4(1-\dot{\rho}^{2J})}{(1-\dot{\rho}^2)^2} \\
&= O(T^{\vartheta}).
\end{aligned}$$

Then, we can see that

$$\frac{1}{T^{3\vartheta}}\left\{\frac{1}{T}\sum_{t=K+1}^{T}\dot{\eta}_t^2\dot{\eta}_{t-K}^2+\sum_{j=1}^{J}\left(1-\frac{j}{J+1}\right)\dot{\rho}^{2j}\frac{1}{T}\sum_{t=j+K+1}^{T}\dot{\eta}_{t-j}^2\dot{\eta}_{t-j-K}^2\right\}=O_p(1) \qquad (62)$$

and the probability limit is positive.

Similar results can be obtained for the second to the fourth terms of (61) by complicated but direct calculations and thus the long-run variance estimator is $O_p(T^{3\vartheta})$, which implies $\hat{\omega}_a = O_p(T^{3\vartheta/2})$. From this result and the convergence order of the numerator given by (54), we obtain the theorem.■

Proof of Theorem 3

As given by Lemma 1, we can apply the B–N decomposition to each $\eta_{i,t}\eta_{i,t-K}$. We can also see from Theorem 1 that $\eta_{i,t}\eta_{i,t-K}$ is the dominate term in $\hat{\eta}_{i,t}^*\hat{\eta}_{i,t-K}^*$while the other terms are negligible and the bias becomes as given in Lemma 3 for each i. The rest we have to show is that the FCLT holds for $\sum_{i=1}^{N}\eta_{i,t}\eta_{i,t-K}$. Note that because $\eta_{i,t}$ is obtained by linear transformations of $\varepsilon_{i,t}$, $\eta_{i,t}$ is independent of $\eta_{j,s}$ for all i,j and $s \neq t$. Thus, we can see that $\sum_{i=1}^{N}\eta_{i,t}\eta_{i,t-K}$ is a martingale difference sequence with respect to the sigma-field constructed from $\eta_{1,t},\eta_{2,t-1},\cdots,\eta_{2,t},\eta_{2,t-1},\cdots,\eta_{N,t},\eta_{N,t-1},\cdots$. Because $G_{i,j}$ for $i=1,\cdots,N$ satisfy Lemma A.1(ii), we can see that the conditions of the FCLT given by Theorem 27.14 of Davidson (1994) are satisfied as in the proof of Theorem 1. We then have the theorem.■

References

[1] Bai, J. and J. L. Carrion-i-Silvestre (2013). Testing panel cointegration with unobservable dynamic common factors that are correlated with the regressors. *Econometrics Journal* 16, 222-249.

[2] Bai, J. and S. Ng (2004). A PANIC attack on unit roots and cointegration. *Econometrica* 72, 1127-1177.

[3] Baltagi, B. H. (2008). *Econometric Analysis of Panel Data*. Wiley, Chichester.

[4] Banerjee, A. and J. L. Carrion-i-Silvestre (2013). Cointegration in panel data with breaks and cross-section dependence. *Journal of Applied Econometrics.* DOI:10.1002/jae.2348.

[5] Banerjee, A., M. Marcellino and C. Osbat (2004). Some cautions on the use of panel methods for integrated series of macro-economic data. *Econometrics Journal* 7, 322-340.

[6] Banerjee, A., M. Marcellino and C. Osbat (2005). Testing for PPP: Should we use panel methods? *Empirical Economics* 30, 77-91.

[7] Banerjee, A. and M. Wagner (2009). Panel methods to test for unit roots and cointegration, in T. C. Mills and K. Patterson ed., *Palgrave Handbook of Econometrics, Vol.2*. Palgrave Macmillan, New York.

[8] Breitung, J. and S. Das (2008). Testing for unit roots in panels with a factor structure. *Econometric Theory* 24, 88-108.

[9] Breitung, J. and M. H. Pesaran (2008). Unit roots and cointegration in panels. In L. Matyas and P. Sevestre (eds), *The econometrics of panel data: Fundamentals and recent developments in theory and practice (3rd edition)*. Springer-Verlag, Berlin.

[10] Brillinger, D. R. (1981). *Time Series Data Analysis and Theory.* Holden-Day, San Francisco.

[11] Chang, Y. (2002). Nonlinear IV unit root tests in panels with cross-sectional dependency. *Journal of Econometrics* 110, 261-92.

[12] Chang, Y. (2004). Bootstrap unit root tests in panels with cross-sectional dependency. *Journal of Econometrics* 120, 263-293.

[13] Chang, Y. and C. M. Nguyen (2012). Residual based tests for cointegration in dependent panels. *Journal of Econometrics* 167, 504-520.

[14] Chang, Y. and W. Song (2009). Unit root tests for panels in the presence of short-run and long-run dependencies: Nonlinear IV approach with fixed N and large T. *Review of Economic Studies* 73, 903-935.

[15] Choi, I. and E. Kurozumi (2012). Model selection criteria for the leads-and-lags cointegrating regression. *Journal of Econometrics* 169, 224-238.

[16] Chortareas, G. and G. Kapetanios (2009). Getting PPP right: Identifying mean-reverting real exchange rates in panels. *Journal of Banking and Finance* 33, 390-404.

[17] Davidson, J. (1994). *Stochastic Limit Theory*. Oxford University Press, Oxford.

[18] De Silva, S., K. Hadri and A. R. Tremayne (2009). Panel unit root tests in the presence of cross-sectional dependence: Finite sample performance and an application. *Econometrics Journal* 12 (2), 340-366

[19] Driscoll, J. C. and A. C. Kraay (1998). Consistent covariance estimation with spatially dependent panel data. *Review of Economics and Statistics* 80, 549-560.

[20] Engle, R. F. and C. W. J. Granger (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica* 55, 251-276.

[21] Frankel, J. A. and A. K. Rose (1996). A panel project on purchasing power parity: Mean reversion within and between countries. *Journal of International Economics* 40, 209-224.

[22] Gengenbach, C., F. C. Palm and J. -P. Urbain (2006). Cointegration testing in panels with common factors. *Oxford Bulletin of Economics and Statistics* 68, 683-719.

[23] Gengenbach, C., F. C. Palm and J. -P. Urbain (2010). Panel unit root tests in the presence of cross-sectional dependencies: Comparison and implications for modelling. *Econometric Reviews* 29, 111-145.

[24] Giraitis, L. and P. C. B. Phillips (2006). Uniform Limit Theory for Stationary Autoregression. *Journal of Time Series Analysis* 27, 51-60.

[25] Granger, C. W. J. and P. Newbold (1974). Spurious regression in econometrics. *Journal of Econometrics* 2, 111-120.

[26] Groen, J. J. J. and F. Kleibergen (2003). Likelihood-based cointegration analysis in panels of vector error-correction models. *Journal of Business and Economic Statistics* 21, 295-318.

[27] Hadri, K. and E. Kurozumi (2008). A simple panel stationarity test in the presence of cross-sectional dependence. *Global COE Hi-Stat discussion paper series 16.*

[28] Hadri, K. and E. Kurozumi (2012). A simple panel stationarity test in the presence of serial correlation and a common factor. *Economics Letters* 115 (1)

[29] Hadri, K. and Y. Rao (2008). Panel stationarity test with structural breaks. *Oxford Bulletin of Economics and Statistics* 70 (2), 245-269. , 31-34.

[30] Hanck, C. (2009). Cross-sectional correlation robust tests for panel cointegration. *Journal of Applied Statistics* 36, 817-833.

[31] Hansen, B. E. (1992). Tests for parameter instability in regressions with I(1) processes. *Journal of Business and Economic Statistics* 10, 321-335.

[32] Harris, D., B. P. McCabe and S. Leybourne (2003). Some limit theory for autocovariances whose order depends on sample size. *Econometric Theory* 19, 829-864.

[33] Harris, D., S. Leybourne and B. P. McCabe (2005). Panel stationarity tests for purchasing power parity with cross-sectional dependence. *Journal of Business and Economic Statistics* 23, 395-409.

[34] Hayakawa, K. and E. Kurozumi (2008). The role of "leads" in the dynamic OLS estimation of cointegrating regression models. *Mathematics and Computers in Simulation* 79, 555-560.

[35] Jansson, M. (2005). Point optimal tests of the null hypothesis of cointegration. *Journal of Econometrics* 124, 187-201.

[36] Kao, C. (1999). Spurious regression and residual-based tests for cointegration in panel data. *Journal of Econometrics* 90, 1-44.

[37] Kejriwal, M. and P. Perron (2008). Data dependent rules for selection of the number of leads and lags in the dynamic OLS cointegrating regression. *Econometric Theory* 24, 1425-1441.

[38] Kurozumi, E. and Y. Arai (2008). Test for the null hypothesis of cointegration with reduced size distortion. *Journal of Time Series Analysis* 29, 476-500.

[39] Kurozumi, E. and K. Hayakawa (2009). Asymptotic properties of the efficient estimators for Cointegrating regression models with serially dependent errors. *Journal of Econometrics* 149, 118-135.

[40] Kwiatkowski, D., P. C. B. Phillips, P. Schmidt and Y. Shin (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics* 54, 159-178.

[41] Maddala, G. S. and S. W. Wu (1999). A comparative study of unit root tests with panel data and a new simple test. *Oxford Bulletin of Economics and Statistics* 61, 631-652.

[42] Magdalinos, T. and P. C. B. Phillips (2009). Limit theory for cointegrated systems with moderately integrated and moderately explosive regressors. *Econometric Theory* 25, 482-526.

[43] Matsuki, T. (2015). Linear and nonlinear comovement in southeast Asian local currency bond markets: A stepwise multiple testing approach. *Empirical Economics*, forthcoming.

[44] McCoskey, S. and C. Kao (1999). A residual-based test for the null of cointegration in panel data. *Econometric Reviews* 17, 57-84.

[45] Moon, H. R. and B. Perron (2004). Testing for a unit root in panels with dynamic factors. *Journal of Econometrics* 122, 81-126.

[46] Moon, H. R. and B. Perron (2012). Beyond panel unit root tests: Using multiple testing to determine the nonstationarity properties of individual series in a panel . *Journal of Econometrics* 169, 29-33.

[47] Newey, W. K. and K. D. West (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55, 703-708.

[48] Ng, S. (2008). A simple test for non-stationarity in mixed panels. *Journal of Business and Economic Statistics* 26, 113-127.

[49] O'Connell, P. G. J. (1998). The overvaluation of purchasing power parity. *Journal of International Economics* 44, 1-19.

[50] Palm, F. C., S. Smeekes and J. -P. Urbain (2011). Cross-sectional dependence robust block bootstrap panel unit root tests. *Journal of Econometrics* 163, 85-104.

[51] Papell, D. H. (1997). Searching for stationarity: Purchasing power parity under the current float. *Journal of International Economics* 43, 313-332.

[52] Pedroni, P. (1999). Critical values for cointegration Tests in heterogeneous panels with multiple regressors. *Oxford Bulletin of Economics and Statistics* 61, 653-670.

[53] Pedroni, P. (2004). Panel cointegration: Asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis. *Econometric Theory* 20, 597-625.

[54] Pedroni, P. L., T. J. Vogelsang, M. Wagner and J. Westerlund (2015). Nonparametric rank tests for non-stationary panels. *Journal of Econometrics* 185, 378-391.

[55] Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-sectional dependence. *Journal of Applied Econometrics* 22, 265-312.

[56] Phillips, P. C. B. and B. E. Hansen (1990). Statistical inference in instrumental variables regression with I(1) processes. *Review of Economic Studies* 57, 99-125.

[57] Phillips, P. C. B. and M. Loretan (1991). Estimating long-run economic equilibria. *Review of Economic Studies* 58, 407-436.

[58] Phillips, P. C. B. and T. Magdalinos (2007a). Limit theory for moderate deviations from a unit root. *Journal of Econometrics* 136, 115-130.

[59] Phillips, P. C. B. and T. Magdalinos (2007b). Limit theory for moderate deviations from a unit root under weak dependence, in G. D. A. Phillips and E. Tzavalis, ed., *The Refinement of Econometric Estimation and Test Procedures: Finite Sample and Asymptotic Analysis*. Cambridge University Press, Cambridge.

[60] Phillips, P. C. B. and H. R. Moon (1999). Linear regression limit theory for nonstationary panel data. *Econometrica* 67, 1057-1111.

[61] Quintos, C. E. and P. C. B. Phillips (1993). Parameter constancy in cointegrating regressions. *Empirical Economics* 18, 675-706.

[62] Saikkonen, P. (1991). Asymptotically efficient estimation of cointegration regressions. *Econometric Theory* 7, 1-21.

[63] Shin, Y. (1994). A residual-based test of the null of cointegration against the alternative of no cointegration. *Econometric Theory* 10, 91-115.

[64] Stock, J. H. and M. W. Watson (1993). A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica* 61, 783-820.

[65] Wagner, M. (2008a). On PPP, unit roots and panels. *Empirical Economics* 35, 229-249.

[66] Wagner, M. (2008b). The carbon Kuznets curve: A cloudy picture emitted by bad econometrics? *Resource and Energy Economics* 30, 388-408.

[67] Westerlund, J. (2005). A panel CUSUM test for the null of cointegration. *Oxford Bulletin of Economics and Statistics* 67, 231-262.

[68] Westerlund, J. (2008). Panel cointegration tests of the Fisher effect. *Journal of Applied Econometrics* 23, 193-233.

[69] Westerlund, J. and D. L. Edgerton (2007). A panel bootstrap cointegration test. *Economics Letters* 97, 185-190.

[70] Westerlund, J. and D. L. Edgerton (2008). A simple test for cointegration in dependent panels with structural breaks. *Oxford Bulletin of Economics and Statistics* 70, 665-704.

[71] Yabe, R. (2012). Limiting Distribution of the Score Statistic under Moderate Deviation from a Unit Root in MA(1). *Journal of Time Series Analysis* 33, 533-541.